

**LOW POWER AND RELIABLE  
DESIGN METHODOLOGIES FOR 3D ICS**

A Dissertation  
Presented to  
The Academic Faculty

by

Moongon Jung

In Partial Fulfillment  
of the Requirements for the Degree  
Doctor of Philosophy in the  
School of Electrical and Computer Engineering

Georgia Institute of Technology  
May 2014

Copyright © Moongon Jung 2014

**LOW POWER AND RELIABLE  
DESIGN METHODOLOGIES FOR 3D ICS**

Approved by:

Dr. Sung Kyu Lim, Advisor  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Madhavan Swaminathan  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Saibal Mukhopadhyay  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Azad Naeemi  
School of Electrical and Computer  
Engineering  
*Georgia Institute of Technology*

Dr. Suresh K. Sitaraman  
School of Mechanical Engineering  
*Georgia Institute of Technology*

Date Approved: February 28, 2014

*Dedicated to my family  
for their endless love and support.*

## ACKNOWLEDGEMENTS

There were ups and downs throughout my doctoral study. Without the help from many people, I could not finish this five-year-long journey successfully.

My first thanks go to my advisor, Prof. Sung Kyu Lim, for his professional and insightful guidance on my research. He gave me the chance to study 3D ICs at Georgia Tech Computer-Aided Design (GTCAD) laboratory, a leading research group in the 3D IC field. After joining his group, a lot of changes happened in my life in a positive way. I hope he continues his success with great people.

I would like to thank Prof. Madhavan Swaminathan and Prof. Saibal Mukhopadhyay for their insightful suggestions and delightful comments on my research. Also, I would like to thank Prof. Azad Naeemi and Prof. Suresh K. Sitaraman for serving as my dissertation committee members.

I would like to extend my thanks to all my colleagues. All previous and current GTCAD group members: Dr. Michael Healy, Mohit Pathak, Dr. Dae Hyun Kim, Dr. Xin Zhao, Dr. Krit Athikulwongse, Dr. Young-Joon Lee, Chang Liu, Taigon Song, Shreepad Panth, Hemant Sane, Dr. Daniel Limbrick, Woongrae Kim, Yarui Peng, Sandeep Samal, Yang Wan, Steven Zhang, Taesik Na, and Neela Lohith.

I had a wonderful chance to work with Prof. David Z. Pan at the University of Texas at Austin and Prof. Yiyu Shi at the Missouri University of Science and Technology. They are undoubtedly world leaders of research and education. Also, I am thankful for the colleagues who worked with me remotely, Joydeep Mitra, Jiwoo Park, and Chun Zhang for insightful discussions and valuable comments.

During the project with Intel, I met the following people online and offline whom I am thankful for: Dr. Paul Fischer, Dr. Hong Wang, Dr. Patrick Morrow, Dr. Bob Rao, Dr. Greg Taylor, Clair Webb, Dr. Vijay Pitchumani, Dr. Debabrata Mohapatra, Dr. Devangkumar Jariwala, and Dr. Nilanjan Banerjee.



My sincere thanks go to my parents and parents-in-law for their endless devotion and support. Without them, this research as well as myself would not even exist. I would like to thank my grandmother for her bottomless love throughout my life. My only brother, Jihun, has been taking over my responsibility ever since I started studying abroad. With his help, I have been able to study without worrying about family. In addition, I would like to express my appreciation to all the relatives for supporting my family throughout this Ph.D. study.

Sooyeon, my lovely daughter, your presence itself is the happiness and joy. No words can describe this. Last, but not least, nothing was possible without my lovely wife, the best person that I ever met and will meet, Yeohun Kim. You made me a better person, and I will keep trying to become a better person for all of us. I am confident that we together will make a wonderful life, and when we look back on our lives in the future, we would say this was a happy life.

## TABLE OF CONTENTS

ACKNOWLEDGEMENTS . . . . .	iv
LIST OF TABLES . . . . .	xii
LIST OF FIGURES . . . . .	xv
LIST OF SYMBOLS OR ABBREVIATIONS . . . . .	xxiii
SUMMARY . . . . .	xxiii
CHAPTER I INTRODUCTION . . . . .	1
1.1 Contributions . . . . .	2
1.2 Organization . . . . .	5
CHAPTER II ORIGIN AND HISTORY OF THE PROBLEM . . . . .	6
2.1 3D Power Distribution Network Analysis and Design . . . . .	6
2.2 Thermo-Mechanical Stress in TSV-based 3D ICs . . . . .	6
2.3 Low Power 3D IC Designs . . . . .	7
CHAPTER III POWER-DELIVERY NETWORK ANALYSIS AND OPTIMIZATION FOR 3D ICS . . . . .	9
3.1 Backgrounds . . . . .	9
3.1.1 Existing Works . . . . .	9
3.1.2 P/G TSV Impacts on 3D IC Layouts . . . . .	10
3.2 Non-regular Power/Ground TSV Placement Algorithm . . . . .	13
3.2.1 Equivalent Circuit for Series Resistors . . . . .	13
3.2.2 Non-regular P/G TSV Placement Algorithm . . . . .	15
3.3 Experimental Results of Non-Regular P/G TSV Placement . . . . .	16
3.4 TSV RC Variation . . . . .	17
3.4.1 $R_{TSV}$ variation . . . . .	18
3.4.2 $C_{TSV}$ variation . . . . .	19
3.5 3D Power Supply Noise Analysis with TSV RC Variation . . . . .	20
3.6 Experimental Results with TSV RC Variations . . . . .	20
3.6.1 Impact of TSV RC Variation Range . . . . .	21
3.6.2 Impact of Number of Variation Sources . . . . .	22

3.6.3	Impact of Number of C4 Bumps . . . . .	22
3.6.4	Impact of TSV Size . . . . .	23
3.7	Summary . . . . .	24
CHAPTER IV TSV STRESS-AWARE FULL-CHIP MECHANICAL RELIABILITY ANALYSIS AND OPTIMIZATION FOR 3D ICS . . . . .		25
4.1	Detailed Baseline Modeling . . . . .	25
4.1.1	3D FEA Simulation . . . . .	27
4.1.2	Impact of TSV Liner and Landing Pad . . . . .	28
4.1.3	Impact of Cu Diffusion Barrier . . . . .	29
4.1.4	Stress Influence Zone . . . . .	31
4.1.5	Anisotropic Material Property of Silicon . . . . .	32
4.2	Full-Chip Stress and Reliability Analysis . . . . .	34
4.2.1	Linear Superposition Principle . . . . .	34
4.2.2	Stress Analysis with Multiple TSVs . . . . .	34
4.2.3	Mechanical Reliability Analysis . . . . .	35
4.2.4	Validation of Linear Superposition Method . . . . .	36
4.2.5	Limit of Linear Superposition Method . . . . .	39
4.2.6	Full-Chip Analysis Flow . . . . .	41
4.2.7	Scalability of Algorithm . . . . .	42
4.3	Full-Chip Simulation Results . . . . .	44
4.3.1	Overall Comparison . . . . .	45
4.3.2	Impact of TSV pitch . . . . .	46
4.3.3	Impact of TSV Size . . . . .	47
4.3.4	Impact of Liner Thickness . . . . .	48
4.3.5	Reliability of Block-Level 3D Design . . . . .	49
4.3.6	Impact of TSV Re-placement . . . . .	50
4.4	TSV and Device-level Variation Study . . . . .	51
4.4.1	Simulation Setup . . . . .	52
4.4.2	Variation Impact on Stress & Reliability . . . . .	53
4.4.3	Variation Impact on Device Mobility . . . . .	55
4.4.4	Variation Impact on Cell Delay . . . . .	57

4.5	Full Chip-level Variation Study . . . . .	57
4.5.1	Challenges . . . . .	57
4.5.2	Divide-and-Conquer Approach . . . . .	58
4.5.3	Quality of Stress Models . . . . .	60
4.5.4	Compact RSM Model vs. Lookup Table . . . . .	60
4.5.5	Multiple TSVs under Variations . . . . .	60
4.6	Variation-Tolerant Full-Chip Design Methods . . . . .	64
4.6.1	Simulation Settings . . . . .	64
4.6.2	Robust Design with TSV Pitch Adjustment . . . . .	65
4.6.3	Robust Design with TSV KOZ Size Adjustment . . . . .	66
4.6.4	Tungsten vs. Copper TSV on Full-Chip Robustness . . . . .	68
4.6.5	Design Style Choice: Block-level vs Gate-level 3D ICs . . . . .	68
4.7	Summary . . . . .	70
CHAPTER V CHIP/PACKAGE MECHANICAL STRESS IMPACT ON 3D IC RE-		
LIABILITY AND MOBILITY VARIATIONS . . . . .		72
5.1	Motivation . . . . .	73
5.2	3D IC/Package Stress Modeling . . . . .	75
5.2.1	Chip/Package Co-Simulation Structure . . . . .	76
5.2.2	Impact of Die Stacking . . . . .	77
5.2.3	Impact of TSV and Bump Alignment . . . . .	78
5.3	Mobility Variation Modeling . . . . .	79
5.3.1	Need for True 3D Chip/Package Stress Model . . . . .	79
5.3.2	Piezoresistivity . . . . .	80
5.3.3	Mobility Variation: 2D vs. 3D Stress . . . . .	81
5.4	Handling Full-Chip/Package: The LVLS Method . . . . .	84
5.4.1	Lateral and Vertical Linear Superposition . . . . .	84
5.4.2	Full-Chip/Package Stress Analysis Flow . . . . .	86
5.4.3	Validation of LVLS . . . . .	87
5.5	Full-Stack Timing Variation Analysis . . . . .	89
5.5.1	Full-Stack Device Mobility Variation . . . . .	89
5.5.2	Chip/Package Stress-Aware Timing Analysis . . . . .	91

5.6	Full-Stack Reliability Analysis Results . . . . .	93
5.6.1	Impact of Package-Bump and $\mu$ -Bump . . . . .	93
5.6.2	Case Study I: Wide-I/O DRAM . . . . .	94
5.6.3	Case Study II: Block-Level 3D IC . . . . .	95
5.7	Full-Stack Timing Analysis Results . . . . .	97
5.7.1	2D vs. 3D Stress Impact on Mobility and Timing . . . . .	97
5.7.2	Case Studies: Block-level and Wide-I/O Style 3D Designs . . . . .	100
5.8	Summary . . . . .	102
CHAPTER VI CRACK ANALYSIS AND OPTIMIZATION FOR THROUGH-SILICON-VIA-BASED 3D ICS . . . . .		104
6.1	TSV Interfacial Crack Modeling . . . . .	105
6.1.1	Energy Release Rate . . . . .	106
6.1.2	3D FEA Simulation . . . . .	107
6.1.3	Impact of TSV Liner and Landing Pad . . . . .	107
6.1.4	Impact of Pitch and Angle among TSVs . . . . .	109
6.1.5	Relative Importance of Pitch over Angle . . . . .	111
6.2	DOE and RSM based Full-Chip TSV Interfacial Crack Modeling . . . . .	112
6.2.1	Designing Experiments . . . . .	113
6.2.2	ERR Model for Regular TSV Placement . . . . .	114
6.2.3	ERR Model for Irregular TSV Placement . . . . .	117
6.2.4	Quality of ERR Model . . . . .	118
6.2.5	Full-chip Analysis Flow . . . . .	119
6.3	Full-Chip TSV Interfacial Crack Simulation Results . . . . .	119
6.3.1	Impact of KOZ . . . . .	120
6.3.2	Impact of TSV Placement Style . . . . .	121
6.3.3	Impact of Liner . . . . .	121
6.3.4	Reliability of Block-Level 3D Design . . . . .	122
6.4	Substrate Crack Modeling . . . . .	124
6.4.1	Crack Propagation Modeling . . . . .	124
6.4.2	Isolated TSV Case . . . . .	124
6.4.3	Single-Aggressor Case . . . . .	128

6.4.4	Two-Aggressor Case . . . . .	130
6.4.5	TSV-Array Case . . . . .	130
6.5	Conclusions . . . . .	133
CHAPTER VII LOW POWER DESIGN METHODOLOGIES FOR 3D ICS . . .		134
7.1	Preliminaries . . . . .	135
7.1.1	Die Stacking Technology . . . . .	135
7.1.2	3D IC Design Flow . . . . .	135
7.2	A Case Study with OpenSPARC T2 Core . . . . .	136
7.2.1	3D Floorplanning Benefits . . . . .	136
7.2.2	Judicious Metal Layer Usage . . . . .	139
7.2.3	Dual-Vth Benefits for 3D ICs . . . . .	141
7.2.4	Folding Functional Unit Blocks . . . . .	143
7.3	Low Power Computing Multi-core 3D Processors . . . . .	145
7.3.1	Baseline Target System Design . . . . .	145
7.3.2	3D Clock Tree Router . . . . .	148
7.4	Block Folding Benefits . . . . .	149
7.4.1	Folding CCX Block . . . . .	151
7.4.2	Folding L2D Block . . . . .	152
7.4.3	Folding L2T Block . . . . .	153
7.4.4	Folding RTX Block . . . . .	155
7.4.5	Second-level Folding SPC Block . . . . .	155
7.5	Face-to-Face Bonding Benefit . . . . .	155
7.5.1	F2F Via Placer . . . . .	156
7.5.2	F2F Impact on Block Folding . . . . .	157
7.6	Full-Chip Power Benefit With Folded Blocks . . . . .	158
7.6.1	3D Pin Partition . . . . .	159
7.6.2	3D Floorplan with Folded Blocks . . . . .	160
7.6.3	Bonding Style Impact: F2B vs. F2F . . . . .	161
7.6.4	Full-chip Design Comparison with Dual-Vth Cells . . . . .	162
7.7	Summary . . . . .	164

CHAPTER VIII CONCLUSIONS . . . . .	166
REFERENCES . . . . .	170
PUBLICATIONS . . . . .	175
VITA . . . . .	179

## LIST OF TABLES

Table 1	Benchmark Circuits. . . . .	17
Table 2	Results of the regular P/G TSV placement. . . . .	17
Table 3	Results of the non-regular P/G TSV placement algorithm. Numbers in parentheses are compared with the regular TSV placement case. . . . .	18
Table 4	Benchmark circuits for TSV RC variation impact study. . . . .	21
Table 5	Impact of TSV RC variation range. . . . .	22
Table 6	Impact of number of variation sources. . . . .	22
Table 7	Impact of number of P/G C4 bumps. . . . .	23
Table 8	Impact of TSV size. Aspect ratio is 10 for all TSVs. . . . .	24
Table 9	Impact of parasitics of nano-scale TSVs. . . . .	24
Table 10	Material properties. . . . .	27
Table 11	Effect of Young's modulus of barrier material on von Mises stress at TSV/barrier interface. Young's modulus of Ta changes while CTE and Poisson's ratio remain same. . . . .	31
Table 12	Von Mises stress comparison between FEA simulations and linear superposition method. . . . .	37
Table 13	Error in Maximum von Mises stress in MPa (and % error) at TSV edge between FEA simulation and linear superposition for both isotropic and anisotropic Si. . . . .	39
Table 14	Maximum error (in MPa) in von Mises stress between FEA simulation and linear superposition. . . . .	41
Table 15	Scalability to full-chip analysis. . . . .	43
Table 16	Benchmark circuits. . . . .	44
Table 17	Impact of TSV size on the maximum von Mises stress. The numbers in parentheses are % reduction compared to TSV large case. . . . .	48
Table 18	Impact of liner thickness on the number of TSVs violating von Mises criterion. The numbers in parentheses are % reduction compared to the 125 nm thick liner case. . . . .	49
Table 19	Comparison between gate-level and block-level design. . . . .	50
Table 20	Impact of TSV re-placement on von Mises stress distribution, wirelength, and longest path delay. . . . .	51
Table 21	Piezo-resistivity coefficient ( $TPa^{-1}$ ) in (100) Si wafer [1]. . . . .	55
Table 22	Quality of the compact stress models for Cu TSV. . . . .	61



Table 23	Comparison between our RSM-model-based analysis and lookup-table-based analysis [2]. TSV pitch is $10\mu m$ . The numbers in parenthesis are % reduction in runtime compared with the lookup table case. . . . .	62
Table 24	Von Mises stress comparison between FEA and the proposed method. Error = ours - FEA. . . . .	64
Table 25	Impact of TSV pitch on full-chip reliability and performance under material property variations. TSV material is Cu for all cases, and $\sigma_{CTE}$ and $\sigma_Y$ of each material are set as 10% of the nominal values. . . . .	67
Table 26	Maximum von Mises stress and longest path delay from 1000 Monte Carlo simulations. Numbers in parenthesis are % increase compared with the <i>no variation</i> ( $\sigma = 0$ ) case (= [2]). . . . .	70
Table 27	Piezoresistive coefficient ( $TPa^{-1}$ ) in (100) Si wafer [3]. . . . .	81
Table 28	Von Mises stress comparison between FEA and LVLS for a 4-die stack structure (die0). Error = LVLS - FEA. (At TSV edge, typical von Mises stress level is around 900MPa.) . . . . .	88
Table 29	Reliability in wide-I/O DRAM. . . . .	95
Table 30	Mechanical reliability in block-level 3D IC. TSV blocks are shown in Figure 55. . . . .	96
Table 31	Benchmark circuits with TSV KOZ = $1.0\mu m$ . . . . .	97
Table 32	Block-level and wide-I/O style 3D IC designs. Package-bump pitch is $200\mu m$ . . . . .	101
Table 33	Quality of ERR model. . . . .	118
Table 34	Validation of ERR model. Simulation case shows pitch for regular TSV placement and number of aggressors for irregular TSV placement. . . . .	119
Table 35	Benchmark circuits. . . . .	120
Table 36	Comparison between gate-level and block-level designs. . . . .	124
Table 37	3D interconnect settings. . . . .	136
Table 38	Comparison between 2D and 3D designs with a target clock period of 1.5ns. Numbers in parentheses are (intra block/inter block) breakdown. . . . .	138
Table 39	Impact of metal layer usage in intra-block level routing on power consumption for 2D and 3D designs. Power is normalized to the case of intra-block routing up to M5. . . . .	140
Table 40	Impact of intra-block metal layer usage on intra-block and inter-block design metrics in the 3D design. Target clock period is 1.5ns and numbers in parentheses are difference with respect to the case of intra-block routing up to M5. . . . .	141

Table 41	Comparison between 2D (intra-block routing up to M5), 3D without FUB folding (intra-block routing up to M7), and 3D with FUB folding designs with a target clock period of 1.5ns. Dual-Vth design technique is applied to all cases. Numbers in parentheses are difference against the 2D design.	145
Table 42	Comparison between 2D and 3D block-level designs with a target clock frequency of 500MHz. Numbers in parentheses are differences against the 2D design. . . . .	147
Table 43	Clock tree comparison: 2D vs. 3D (core/cache). . . . .	149
Table 44	2D design characteristics used for block folding candidate selection. Long wires are defined as wires longer than 100X standard cell height. CPU clock runs at 500MHz and I/O clock at 250MHz. . . . .	150
Table 45	Wirelength comparison between 2D and 3D CCX. . . . .	152
Table 46	Comparison between 2D and 3D L2D designs. . . . .	153
Table 47	L2T die partitioning schemes. . . . .	153
Table 48	Comparison between 2D, 3D without block folding (core/cache, F2B), and 3D with block folding (5 types of blocks folded, F2F) designs. Dual-Vth design technique is applied to all cases. Numbers in parentheses are difference against the 2D excluding HVT cell count which shows % of total cell count. . . . .	163
Table 49	Intra-block and inter-block level comparisons between 2D and 3D designs shown in Figure 100. Dual-Vth design technique is applied to all cases. $\Delta$ is the difference against the 2D design. . . . .	165

## LIST OF FIGURES

Figure 1	Target 3D structure with via-first TSVs. P/G TSVs are vertically connected with stacked local via arrays. . . . .	11
Figure 2	Routing congestion in M5 and M6 due to regularly placed P/G TSVs . . . . .	11
Figure 3	Power distribution network layout. Additional routing resources are needed for P/G net to detour P/G TSVs and signal TSVs. . . . .	12
Figure 4	Resistive P/G network. (a) Series resistor chain in P/G network. (b) Series resistor equivalent circuit. . . . .	13
Figure 5	Boundary voltage and current change due to P/G TSV insertion. . . . .	15
Figure 6	Flow chart of the non-regular P/G TSV placement algorithm. . . . .	16
Figure 7	TSV RC variations caused by TSV dimension variations. (a) TSV resistance vs. TSV diameter and height variation. (b) TSV capacitance vs. TSV diameter and oxide thickness variation. . . . .	19
Figure 8	Layout and power supply noise maps in the bottom die of ind1. (a) Layout. White rectangles are P/G TSVs and green rectangles are signal TSVs. (b) Dynamic noise map. Red color represents noisy spots. (c) Static noise map. . . . .	21
Figure 9	Baseline TSV structure. (a) TSV <sub>A</sub> cell occupying three standard cell rows (KOZ = 1.205 $\mu m$ ). (b) TSV <sub>B</sub> cell occupying four standard cell rows (KOZ = 2.44 $\mu m$ ). . . . .	26
Figure 10	Effect of TSV structures on normal stress components. (a) $\sigma_{rr}$ stress. (b) $\sigma_{\theta\theta}$ stress. . . . .	28
Figure 11	Effect of liner material/thickness on $\sigma_{rr}$ stress. . . . .	29
Figure 12	Effect of Cu diffusion barrier on stress. (a) $\sigma_{\theta\theta}$ stress. (b) Von Mises stress. . . . .	30
Figure 13	Effect of TSV size on stress influence zone. (a) $\sigma_{rr}$ stress from TSV edge. (b) $\sigma_{rr}$ stress along normalized distance (distance/TSV diameter). . . . .	32
Figure 14	Effect of anisotropic Si material property on stress (a) $\sigma_{xx}$ stress. (b) $\sigma_{zz}$ stress. . . . .	33
Figure 15	Sample stress comparison between FEA simulation and linear superposition method. (a) FEA result ( $\sigma_{xx}$ ). (b) Ours ( $\sigma_{xx}$ ). (c) FEA vs. ours ( $\sigma_{xx}$ ) along the white line in (a). (d) FEA result (von Mises stress). (e) Ours (von Mises stress). (f) FEA vs. ours (von Mises stress) along the white line in (d). . . . .	38
Figure 16	Von Mises stress comparison between isotropic and anisotropic Si. Stress is taken along the white line in Figure 15(d). . . . .	39
Figure 17	Simulation structures for limit of linear superposition method. (a) Two TSVs. (b) Three TSVs. . . . .	40

Figure 18	Von Mises stress along the center line in Figure 17(a) with two TSVs with different pitches. (a) Two TSVs with pitch of $7.5\mu m$ . (b) Two TSVs with pitch of $25\mu m$ . . . . .	41
Figure 19	Simulation flow . . . . .	42
Figure 20	Impact of TSV structure, TSV placement style, and KOZ size on the maximum von Mises stress. (a) Designs with TSV <sub>A</sub> cell (KOZ = $1.205\mu m$ ). (b) TSV <sub>B</sub> cell (KOZ = $2.44\mu m$ ). . . . .	45
Figure 21	Close-up shots of layouts and von Mises stress maps. (a) Irreg <sub>A</sub> . (b) Reg <sub>A</sub> . (c) Von Mises stress map of Irreg <sub>A</sub> . (d) Von Mises stress map of Reg <sub>A</sub> . . . . .	46
Figure 22	Impact of TSV pitch on maximum von Mises stress. . . . .	47
Figure 23	Impact of liner thickness on the maximum von Mises stress of circuits with TSV <sub>A</sub> cell. . . . .	48
Figure 24	Layout of block-level design (TSV pitch = $15\mu m$ ). White rectangles are TSV landing pads. (a) Full-chip layout. (b) Close-up shot of the red box in (a). . . . .	49
Figure 25	TSV re-placement to reduce von Mises stress. TSV landing pads are white rectangles. (a) Original layout. (b) After TSV re-placement. . . .	50
Figure 26	Significant Cu Young's modulus variation due to the variations on grain size and orientation in Cu TSV (focused ion beam (FIB) image with gray contrast [4]). . . . .	52
Figure 27	Cu CTE variation impact on $\sigma_{xx}$ stress (FEA results). (a) under -15% variation, (b) under +15% variation. . . . .	53
Figure 28	Cu variation impact on $\sigma_{xx}$ under various variation ranges (FEA results). (a) Cu CTE variation, (b) Cu Young's modulus variation. . . . .	54
Figure 29	Cu CTE variation impact on carrier mobility. (a) hole mobility map (Cu CTE: -15%). (b) electron mobility map (Cu CTE: -15%). (c) hole mobility map (Cu CTE: +15%). (d) electron mobility map (Cu CTE: +15%). . . . .	56
Figure 30	Impact of material property variations on mobility. CTE and Young's modulus of materials in Cu and W TSVs vary from -30% to +30%. . . .	57
Figure 31	Cell FO4 delay variations under material property variations. 1000 Monte Carlo simulations are performed and TSV material is Cu. NAND_X1 is placed along the X-axis from TSV center. Delay is normalized to <i>no variation</i> case. (a) Rise delay distribution. (b) Fall delay distribution. . . . .	58
Figure 32	Stress comparison between FEA, lookup-table-based analysis [2], and RSM-model-based analysis near TSV edge. . . . .	62

Figure 33	Von Mises stress comparison between FEA simulation and our method. TSV pitch is $10\mu m$ . (a) FEA result. (b) Ours. (c) FEA vs. ours along the white line in (a). . . . .	64
Figure 34	Layouts and thermal maps of different 3D IC design styles. Yellow rectangles are TSVs. (a) Gate-level design with regular TSV placement (#TSV: 1024). (b) Thermal map of (a). (c) Block-level design (#TSV: 727). (d) Thermal map of (c). . . . .	65
Figure 35	Impact of KOZ size on full-chip performance under variations. $\sigma_{CTE}$ and $\sigma_Y$ of each material are set as 10% of nominal values. Longest path delay is normalized to <i>no variation</i> case (= [2]). . . . .	68
Figure 36	Impact of TSV material on full-chip reliability and performance under material property variations. $\sigma_{CTE}$ and $\sigma_Y$ of each material are set as 10% of the nominal value. (a) Maximum von Mises stress distribution. (b) Longest path delay (normalized to <i>no variation</i> case [2]) distribution. . . . .	69
Figure 37	Impact of 3D IC design styles on full-chip reliability and performance under variations. (a) Max von Mises stress distribution. (b) Longest path delay (normalized to <i>no variation</i> case [2]) distribution. . . . .	70
Figure 38	Impact of bumps and underfill on the stress of device layer (= red line). (a) TSV only [5] (b) TSV + $\mu$ -bump (c) TSV + package-bump (d) TSV + $\mu$ -bump + package-bump. (e) Deformed structure of (b). (f) Deformed structure of (c). Both (e) and (f) are drawn with 10X deformation scale factor. . . . .	73
Figure 39	Impact of package components on the stress ( $\sigma_{rr}$ ) around TSV on device layer (FEA results). . . . .	74
Figure 40	Comparison of impact of package-bump on the device layer stress ( $\sigma_{rr}$ ) between 2D IC and 3D IC (2-die stack) (FEA results). . . . .	75
Figure 41	Side view of baseline chip/package simulation structures. (a) 2-die stack (b) 4-die stack. . . . .	76
Figure 42	Impact of die stacking on device layer stress (FEA results). (a) $\sigma_{rr}$ stress on device layer in each die in 4-die stack. (b) Von Mises stress in each die in a 4-die stack. . . . .	77
Figure 43	Impact of relative position between TSV/ $\mu$ -bump and package-bump on von Mises stress. (a) Initial position. (b) Final position where TSV/ $\mu$ -bump are shifted by $300\mu m$ from package bump center. (c) Von Mises stress at TSV edge along the distance between TSV/ $\mu$ -bump and package-bump (FEA results). . . . .	79
Figure 44	Mobility variation map around a single TSV. (a) Hole mobility (2D biaxial stress). (b) Electron mobility (2D biaxial stress). (c) Hole mobility in die0 in 4-die stack (3D stress with package components). (d) Electron mobility in die0 in 4-die stack (3D stress with package components). For both (c) and (d) TSV, $\mu$ -bump, and package-bump are vertically aligned. . . . .	83

Figure 45	Mobility variation range of a single TSV with different stress cases. Mobility variation numbers are collected along the x-axis and the y-axis from a TSV center on device layers. (a) Hole mobility under 2D and 3D stress without package components. (b) Electron mobility under 2D and 3D stress without package components. (c) Hole mobility under 3D stress with package components in 4-die stack. (d) Electron mobility under 3D stress with package components in 4-die stack. . . . .	84
Figure 46	Illustration of vertical linear superposition with a 2-die stack structure. Stress is extracted along the red line on device layer from each structure using FEA tool. . . . .	85
Figure 47	Vertical linear superposition of $\sigma_{rr}$ stress in a 2-die stack shown in Figure 46	86
Figure 48	Sample stress comparison between FEA and LVLS. (a) Test structure. (b) Close-up shot of von Mises stress map (using LVLS) taken from the red box in (a) on the device layer in die0 in a 4-die stack. (c) FEA vs. LVLS along the red line in (b). . . . .	89
Figure 49	Normal stress components induced by package-bump on device layers (FEA results). (a) Stress in die0 along the x-direction. (b) Stress in die1 and die2 along the x-direction. (c) Stress in die0 along the y-direction. (d) Stress in die1 and die2 along the y-direction. . . . .	90
Figure 50	Mobility variation map with 441 TSVs/ $\mu$ -bumps (black dots) and 9 C4 bumps (white circles) (LVLS results). (a) Hole mobility variation map in die0. (b) Hole mobility variation map in die2. (c) Electron mobility variation map in die0. (d) Electron mobility variation map in die2. . . .	91
Figure 51	Mobility variation impact on cell FO4 delay. (a) Rise delay dependency on hole mobility variation (INV_X1). (b) Fall delay dependency on electron mobility variation (INV_X1). (c) Rise delay dependency on hole mobility variation (NAND_X1). (d) Fall delay dependency on electron mobility variation (NAND_X1). . . . .	92
Figure 52	Von Mises stress map for TSVs (die0 in a 4-die stack). Colored dots are TSVs and white circles are package-bumps. (a) Test structure. (b) Close-up shot of red box in (a) . . . . .	93
Figure 53	Impact of package components and die stacking on the mechanical reliability of TSVs (900 TSVs in each die). . . . .	94
Figure 54	Mechanical reliability in wide I/O DRAM. 1024 TSVs are placed in the middle of a chip. (a) Package-bumps are placed underneath TSV arrays. (b) Package-bumps are placed 200 $\mu m$ apart from TSV arrays. (not drawn in scale.) . . . . .	95
Figure 55	Mechanical reliability in block-level 3D IC. (a) Sample layout of block-level design. (b) Von Mises stress map for TSVs in red box in (a). . . .	96
Figure 56	Cell mobility variation histogram in die0 in 4-die stack (ckt2). (a) Electron mobility. (b) Hole mobility. . . . .	98

Figure 57	Impact of 2D and 3D stress cases on the longest path delay (LPD) and total negative slack (TNS). Timing numbers are normalized to the no-stress case. TSV KOZ is $1\ \mu m$ for all cases. (a) LPD variation. (b) TNS variation. . . . .	99
Figure 58	Full-chip layout (die0 in 4-die stack) with the highlighted longest path. White squares are TSVs and yellow circles are package bumps. (cell mobility naming convention: e.g., Em8_Hp4 = <u>e</u> lectron mobility <u>minus</u> 8 % and <u>h</u> ole mobility <u>plus</u> 4 %). (a) Layout of ckt2 (KOZ = $1.0\mu m$ ). (b) Cells in red circle in (a). (c) Close-up shot of green circle (1) in (b). (d) Close-up shot of green circle (2) in (b). (e) Layout of ckt2 (KOZ = $0.3\mu m$ ). (f) Cells in red circle in (e). (g) Close-up shot of green circle (3) in (f). (d) Close-up shot of green circle (4) in (f). . . . .	100
Figure 59	Layout and mobility variation map of wide-I/O style design (ckt2.wide). (a) Layout of die0 in 4-die stack with the highlighted cells in the critical path. (b) Close-up shot of red circle in (a). (c) Hole mobility variation map. (d) Electron mobility variation map (LVLS results). . . . .	101
Figure 60	Impact of 2D and 3D stress cases on the longest path delay (LPD) and total negative slack (TNS) in block-level and wide-I/O style 3D IC designs. TSV KOZ is $1.7\mu m$ for all cases. (a) LPD variation. (b) TNS variation. . . . .	102
Figure 61	TSV interfacial crack structure under a negative thermal load. (a) Side view with initial crack length of $d$ . (b) Top view. . . . .	105
Figure 62	Baseline TSV structure. (a) TSV <sub>A</sub> cell occupying four standard cell rows (KOZ = $2.44\mu m$ ). (b) TSV <sub>B</sub> cell (KOZ = $1.205\mu m$ ). . . . .	108
Figure 63	Impact of TSV surrounding structures on ERR. . . . .	108
Figure 64	Impact of TSV liner on ERR. Landing pad width of $6\ \mu m$ is used with each liner case. . . . .	109
Figure 65	Impact of TSV pitch on ERR. . . . .	110
Figure 66	Impact of TSV pitch with liner and landing pad ( $6\mu m$ ) on ERR. . . . .	110
Figure 67	Simulation structure for angular dependency. Distance from victim to all aggressors is $d$ . (a) Two aggressor TSVs. Aggressor 1 is fixed and aggressor 2 rotates. (b) Three aggressor TSVs. Aggressor 1 is fixed and aggressor 2 and 3 rotate. . . . .	111
Figure 68	Impact of angle between victim and two aggressor TSVs on ERR. . . . .	111
Figure 69	DOE based simulation points of two aggressor TSVs. . . . .	112
Figure 70	ERR contour map of pitch and angle. . . . .	112
Figure 71	ERR surface map of pitch and $\theta_1$ in three aggressor TSVs case. . . . .	113
Figure 72	Top view of meshed simulation structures for $5\times 5$ TSV block. Orange circles are TSVs. (a) Array type. (b) Staggered type. . . . .	114

Figure 73	Top view of meshed simulation structure for irregular TSV placement with 21 aggressor TSVs. . . . .	115
Figure 74	ERR vs. pitch for array and staggered type. . . . .	115
Figure 75	Top view of meshed simulation structures for (a) line style, (b) cross style. . . . .	116
Figure 76	ERR vs. pitch for line and cross type. . . . .	116
Figure 77	Predicted ERR using DOE and RSM vs. observed ERR. . . . .	118
Figure 78	Impact of keep-out-zone on ERR. TSV <sub>A</sub> cell (KOZ = 2.44 $\mu m$ ) and TSV <sub>B</sub> cell (KOZ = 1.205 $\mu m$ ) (a) Regular TSV placement. (b) Irregular TSV placement. . . . .	120
Figure 79	Close-up shots of layouts and ERR maps. (a) Irreg <sub>B</sub> . (b) Reg <sub>B</sub> . (c) ERR map of Irreg <sub>B</sub> . (d) ERR map of Reg <sub>B</sub> . . . . .	121
Figure 80	Impact of liner material and thickness on maximum ERR. (a) Irreg <sub>B</sub> vs. Reg <sub>B</sub> . (b) Reg <sub>B</sub> vs. Reg <sub>A</sub> . . . . .	122
Figure 81	Layout of block-level design (TSV pitch = 15 $\mu m$ ). White rectangles are TSV landing pads. (a) full-chip layout. (b) close-up shot of the red box in (a). . . . .	123
Figure 82	Top view of a crack (c) initiated from TSV/liner interface and reaching the $n^+$ doped region in arbitrary direction. . . . .	125
Figure 83	Top view of a crack simulation structure of a single TSV with different crack propagation directions. Initial crack length is 0.5 $\mu m$ from TSV/liner interface. . . . .	126
Figure 84	Stress maps around a crack front in an isolated TSV. (a) $\sigma_{yy}$ map with crack propagating with 0° angle. (b) $\sigma_{xx}$ map with crack propagating with 90° angle. . . . .	127
Figure 85	ERR of the crack propagating along the radial direction of a single TSV. . . . .	127
Figure 86	Simulation structure with one aggressor TSV. A crack grows along the x-direction (radial direction). The aggressor TSV either rotates around a victim TSV or moves away from the victim TSV. . . . .	128
Figure 87	ERR of the crack of a victim TSV with one aggressor at different pitches. Initial crack is 0.5 $\mu m$ (from victim TSV/liner interface to liner/substrate interface). Crack length is swept from the initial crack up to the TSV/liner interface of the aggressor with 10 $\mu m$ pitch. . . . .	128
Figure 88	ERR of the crack of a victim TSV with one aggressor at different angles. TSV pitch is 10 $\mu m$ for all cases. . . . .	129
Figure 89	Crack propagation with two aggressor TSVs. (a) Simulation structure to examine if the crack propagates toward one aggressor or in between aggressors. (b) $\sigma_{yy}$ map of (a). . . . .	130
Figure 90	Stress maps of a 5×5 TSV array. (a) $\sigma_{xx}$ map. (b) $\sigma_{yy}$ map. . . . .	131



Figure 91	ERR of the crack of a victim TSV in different locations of a 5×5 TSV array. side_L and ctr_L overlap. Naming convention: e.g., ctr_L = a crack propagates towards left from the victim TSV in the center of a TSV array.	132
Figure 92	Maximum ERR in different TSV array sizes. Inside array: a crack propagates inside a TSV array. Outside array: a crack propagates away from a TSV array. . . . .	132
Figure 93	Die bonding styles. (a) face-to-back. (b) face-to-face. . . . .	135
Figure 94	2D and 3D placement results. (a) 2D design. (b) 3D design with 2979 TSVs. Cyan dots are core-level buffers. Blue and red rectangles are TSV landing pads at M1 and M9, respectively. White arrows represent major inter-block connections. . . . .	137
Figure 95	Cell size usage (%) comparison between 2D and 3D designs. X0 is the smallest cell size. . . . .	139
Figure 96	2D and 3D inter-block routing results. Intra-block routing uses up to M5, and inter-block routing uses up to M9. (a) 2D design. (b) 3D design. . .	140
Figure 97	Power vs. delay curves and HVT cell usage for 2D (intra-block routing up to M5) and 3D (intra-block routing up to M7) DVT designs. . . . .	142
Figure 98	Wirelength distributions of top four largest FUBs in T2 core. . . . .	143
Figure 99	Placement results of a folded FUB and a 3D block-level design with the folded FUB. (a) folded LSU block (TSV#: 596). (b) 3D design with the folded LSU (TSV#: 2411 (1815+596)). . . . .	144
Figure 100	GDSII layouts of 5 design styles of OpenSPARC T2 (full-chip) we compare: (a) 2D design (9x7.9mm <sup>2</sup> ), (b) core/cache stacking (6x6.4mm <sup>2</sup> , #TSV=3,263), (c) core/core stacking (6x6.4mm <sup>2</sup> , #TSV=7,606), (d) block folding with TSVs (6x6.6mm <sup>2</sup> , #TSV=69,091), (e) block folding with F2F (6x6.6mm <sup>2</sup> , #F2F=112,308). Cyan dots inside blocks are intra-block TSVs or F2F vias. . . . .	146
Figure 101	3D clock tree synthesis flow (3D CTS): (1) Intra-block CTS, (2) Die top CTS. Clock TSVs are clock sources of die top. (3) Die bottom CTS. Die top clock tree information is utilized during die bottom CTS through clock TSVs. . . . .	148
Figure 102	Layouts of 3D clock tree of 3D design with four types of blocks folded (F2B). Yellow circles are clock buffers. (a) Die top clock tree. (b) Die bottom clock tree. (c) Close-up shot of white box in (a). (d) Close-up shot of white box in (b). 3D clock tree from clock source module (die bottom) to one L2B block (die top) is shown in (c) and (d). . . . .	149
Figure 103	CCX 2D and 3D layouts. (a) 2D design. CPX is highlighted with white color. (b) 3D design (# TSV=4). . . . .	151
Figure 104	L2D 2D and 3D layouts. (a) 2D design. (b) 3D design (# TSV: 3,119). Blue and red rectangles are TSV landing pads at M1 and M9, respectively.	152

Figure 105	L2T 2D and 3D layouts. (a) 2D design. (b) 3D design of partition #5 in Table 47 (# TSV: 5,073). The top die consists of small macros and synthesized memory. The bottom die contains large macros and logic cells. The total TSV area is 10%. . . . .	154
Figure 106	Impact of L2T die partitioning on wirelength, buffer count, and power. All numbers are normalized to the 2D. . . . .	154
Figure 107	Second-level folding of a SPARC core. 6 FUBs shown in black text are folded (# F2F via: 10,251). . . . .	156
Figure 108	Finding F2F via locations by 3D net routing. (a) Run 3D placer assuming an ideal 3D interconnect. (b) Create 2D-like 3D design files (Verilog, LEF, and DEF). (c) Route 3D nets and extract F2F via locations. . . . .	156
Figure 109	3D net routing. (a) Layout shot after 3D net routing. (b) Close-up shot showing cells in both dies. (c) Close-up shot of 3D net routing showing F2F vias. . . . .	157
Figure 110	Bonding style impact on 3D placement. Blue rectangles are TSV landing pads at M1 and yellow dots are F2F vias. (a) L2D bottom die. (b) L2T bottom die. . . . .	158
Figure 111	Bonding style impact on power in L2T folding. Numbers in parentheses are number of TSVs/F2F vias. . . . .	159
Figure 112	Block folding impact on 3D full-chip power. (a) F2B bonding case. (b) F2F bonding case. . . . .	161

## SUMMARY

The main objective of this dissertation is to explore and develop computer-aided-design methodologies and optimization techniques for reliability, performance, and power of through-silicon-via-based 3D IC designs. Through-silicon-via (TSV), a vertical interconnect element between dies, is the key enabling technology in 3D ICs. This new design element provides unprecedented design freedom as well as challenges. To maximize benefits and overcome challenges in TSV-based 3D ICs, new analysis methodologies and optimization techniques should be developed. Towards the objective, this dissertation includes five research projects.

The first project is on the analysis and optimization of power delivery network in 3D ICs. Because of the increased power density with reduced footprint area, delivering current to all parts of the 3D stack becomes highly challenging. Moreover, increasing power/ground (P/G) TSV count to reduce the power supply noise can cause severe signal net routing congestions. To mitigate the power supply noise issues in 3D ICs with a minimal number of P/G TSVs, a non-regular P/G TSV insertion algorithm has been developed. In addition, the robustness of 3D power delivery network under TSV RC variation is studied.

The second project is on thermo-mechanical stress and reliability analysis methodologies for full-chip-scale 3D IC designs. Because of the coefficients of thermal expansion (CTE) mismatch between TSV fill material such as copper and silicon substrate, thermo-mechanical stress is induced during fabrication process, which can affect device performance or mechanical failures. Finite element analysis (FEA) methods have been widely used to assess this thermo-mechanical stress and reliability issues in 3D ICs. However, FEA methods are computationally expensive and infeasible for large-scale analysis. In this project, a full-chip TSV thermo-mechanical stress and reliability analysis flow based on the linear superposition method is presented which overcomes the limitation of the FEA method. In addition, design optimization methodologies are explored to reduce mechanical reliability

problems in TSV-based 3D ICs.

The third project is on the chip/package co-analysis of thermo-mechanical stress, reliability, and timing for 3D ICs. Most previous works on the thermo-mechanical stress and reliability of TSV-based 3D ICs have been done separately in either chip or package domain. However, package bumps, underfill, and packaging substrate all add further mechanical stress to the 3D IC mounted above it in a non-trivial way. To accurately assess thermo-mechanical reliability problems and device performance variations in 3D IC/package systems, it is imperative to consider the interplay between the stress caused by the TSVs and the one by these packaging elements simultaneously. In this project, a full-chip/package-scale thermo-mechanical stress co-analysis methodology based on the lateral and vertical linear superposition is presented. In addition, design optimization methodologies to reduce the mechanical reliability problems and the mobility and full-stack timing variations caused by the CTE mismatch among the materials are addressed in full-chip/package scale.

The fourth project is on the modeling, analysis, and optimization of TSV cracks. The TSV-induced stress can drive the interfacial cracking between dielectric liner and silicon substrate or the cohesive cracking in dielectric liner and silicon substrate. These cracks may damage transistors nearby, create conducting paths between TSVs (= short circuit), and cause the entire chip operation failure in the worst case. In this project, a fast and efficient full-chip TSV interfacial crack analysis flow is presented based on design of experiments (DOE) and response surface methodology (RSM). The impact of TSV placement structures on the substrate crack growth patterns is also studied.

The last project is on the low power design methodologies for 3D ICs. Power reduction has been one of the most critical design considerations for IC designers. Although 3D IC designs are believed to reduce power consumption due to shorter wirelength, there has not been thorough studies on how to maximize the 3D power benefit. In this project, physical design techniques that are shown to significantly reduce power consumption in 3D ICs are presented. This includes 3D floorplanning, metal layer usage control, and block folding methodologies. The impact of die bonding style, i.e., face-to-back (F2B) and face-to-face (F2F), on 3D power benefit is also presented.

# CHAPTER I

## INTRODUCTION

A major focus of the semiconductor industry in the last 4-5 decades has been to miniaturize ICs by device and interconnect scaling, which is now around 22nm node. While ITRS still predicts further CMOS scaling, e.g., to around 7nm node by the year of 2020 [6], such scaling will reach fundamental physical limit. In addition, the device scaling will be slowed down as next generation lithography methods (e.g., extreme ultraviolet lithography and electron beam lithography) are being pushed back. Or even before that happens, the economy of scaling will require other means for “more Moore” and “more than Moore” integration.

Due to the increasing power, performance, and financial bottlenecks beyond 32-22nm, industry began to look for alternative solutions. This has led to the active research, development, and deployment of thinned and stacked three-dimensional integrated circuits (3D ICs), initially by wire-bond, later by flip-chip, and recently by Through-Silicon-Via (TSV).

TSV is the key enabling technology in 3D IC. This TSV provides vertical signal, power, and thermal paths between the dies in a stack. With 3D integration technology employing TSVs, both the average and maximum distance between components can be substantially reduced by placing them on different dies, which translates into significant savings in delay, power, and area. Moreover, it enables the integration of heterogeneous devices, such as 28nm for high-speed logic and 130nm for analog, making the entire system more compact and efficient.

However, as multiple dies are stacked vertically in 3D ICs, new problems arise. 3D ICs involve disruptive manufacturing technologies compared with conventional 2D ICs. TSVs cause significant thermo-mechanical stress that may seriously affect performance and reliability of circuits. Delivering power to all parts of the 3D stack is also challenging because of the increased power density. In addition, in-depth studies are required to maximize the performance and power benefit of 3D ICs.

The main objective of this dissertation is to explore and develop computer-aided-design (CAD) methodologies and optimization techniques for reliability, performance, and power of TSV-based 3D IC designs. This new design element (= TSV) provides unprecedented design freedom as well as challenges. To maximize benefits and overcome challenges in TSV-based 3D ICs, new analysis and design methodologies as well as optimization techniques should be developed. The physical design methodologies and optimization techniques for 3D ICs should reflect the technological details of today and future as much as possible. In this dissertation, five projects are presented that partially address the aforementioned problems.

### **1.1 Contributions**

The contributions of this dissertation are summarized as follows.

- **A study on the impact of non-regular P/G TSV placement and TSV RC variation on 3D power delivery network:** It is well known that power delivery is a major reliability concern in 3D ICs. P/G TSVs consume a considerable amount of silicon area as well as routing resources unless designed carefully. In this dissertation, a non-regular P/G TSV placement method is presented to suppress the power supply noise to an acceptable level while minimizing the number of P/G TSVs. The results from the conventional regular (= 2D array style) and the proposed non-regular P/G TSV placement are compared in terms of IR-drop noise, P/G TSV count, and other metrics such as wirelength and footprint area. Next, the robustness of 3D power delivery network under TSV RC variation is studied. For this, TSV RC variation due to process variation is modeled. Both static (IR-drop) and dynamic noise (voltage droop) analysis are performed on GDSII-level 3D IC layouts with this TSV RC variation model. The impacts of number of variation sources (P/G TSV count), number of P/G bumps, and TSV RC variation range on the power supply noise are discussed. The impact of P/G TSV size and its variation the 3D PDN quality is also examined.
- **First full-chip thermo-mechanical stress and reliability analysis method:** Because of the coefficients of thermal expansion (CTE) mismatch between TSV fill

material such as copper and silicon substrate, a large thermo-mechanical stress builds up around a TSV. Since this stress affects device performance and mechanical reliability of the 3D stack, an accurate stress assessment is very important. Conventionally, a finite-element-analysis (FEA) method has been utilized to examine the TSV-induced stress and reliability issues. However, this FEA simulation requires huge computing resources and time even for a small number of TSVs, and thus this is not feasible full-chip-scale stress analysis. To overcome this limitation, a fast and accurate full-chip stress and reliability analysis flow based on the linear superposition of stress tensors is presented. This is the first work that studies the TSV stress impact in full-chip 3D designs. This method is then applied to optimize 3D designs for reliability. Then, this tool flow is extended to handle material property variations in the TSV structure caused by process variations and temperature. Finally, material property variation tolerant design methods are explored.

- **Full-stack chip/package co-analysis of thermo-mechanical stress and its impact on reliability and performance:** The thermo-mechanical stress and reliability issues in TSV-based 3D ICs are not solely originated from TSVs, but from packaging elements such as package bumps, underfill, and packaging substrate. These all generate additional stress to the 3D IC in a non-negligible way. However, most previous works have been done separately in either chip or package domain, and thus the thermo-mechanical stress and reliability in 3D IC/package systems cannot be correctly assessed. Moreover, many prior studies on TSV-stress-aware transistor mobility and performance variation assumed the identical stress distribution around TSVs across the stack as these studies did not consider the impact of packaging elements. In this dissertation, a chip/package co-analysis of thermo-mechanical stress method is presented based on lateral and vertical linear superposition. This is the first work that accurately computes stress distributions across the 3D stack considering both chip and packaging elements within a fraction of runtime of FEA simulations. With this method, full-stack reliability and performance analyses are performed, which are then utilized to provide design optimization guidelines to reduce reliability issues and to

suppress performance variations.

- **Analysis and optimization of TSV interfacial crack and substrate crack:**

The most catastrophic mechanical reliability problem caused by TSV-induced stress is a crack. If there is a small defect such as a void around a TSV, the TSV-induced stress can drive the interfacial cracking between dielectric liner and silicon substrate or the cohesive cracking in dielectric liner and silicon substrate. These cracks can jeopardize the reliability of the devices along the propagation path and hence cause the chip function failure in the worst case. In this work, a full-chip TSV interfacial crack analysis and optimization method are presented. First, the TSV interfacial crack growth behavior is modeled based on the realistic TSV structures. These modeling results are then utilized to build a full-chip TSV interfacial crack analysis method that employs design of experiments (DOE) and response surface methodology (RSM). Next, the impact of design knobs such as TSV placement styles and dielectric liner materials are discussed. Additionally, the substrate crack propagation is modeled under different TSV placement scenarios.

- **Physical design methods for low power 3D ICs:** Low power is widely considered as a key benefit of 3D ICs, yet there have been few thorough design studies on how to maximize power benefits in 3D ICs. In this work, physical design techniques that are shown to significantly reduce power consumption in 3D ICs are presented. These methods exploit different design characteristics such as routing resource demand between 2D and 3D designs, and utilize these to optimize power in 3D designs. Traditionally, the power benefit is obtained by 3D floorplanning of 2D IP blocks, which helps reduce long inter-block wires. To further improve power benefits in 3D ICs on top of the traditional 3D floorplanning, intra-block metal layer usage control and functional block folding methods are discussed. In addition, the impact of die bonding style, i.e., face-to-back (F2B) and face-to-face (F2F), on block folding quality and hence 3D power benefit are presented. In all 3D design cases, timing-closed, full-chip GDSII layouts are built and sign-off iso-performance power comparisons with 2D



IC designs are performed. Important design metrics such as area, wirelength, buffer count, timing, and power consumption are compared for 2D and 3D designs.

## ***1.2 Organization***

The rest of this dissertation is organized as follows:

- In Chapter 2, the origin of the problems and the related works are discussed.
- In Chapter 3, the robustness of power delivery network under different P/G TSV placement styles and TSV RC variation are presented.
- In Chapter 4, the full-chip TSV thermo-mechanical stress and reliability analysis flow is presented.
- In Chapter 5, the full-chip/package-scale thermo-mechanical stress, reliability, and performance study is presented.
- In Chapter 6, the TSV interfacial crack and substrate crack analysis and optimization are presented.
- In Chapter 7, the low power design methodologies for 3D ICs are presented.
- In Chapter 8, the conclusions of this dissertation are mentioned, as well as the remarks on the covered topics and possible future works.

## CHAPTER II

### ORIGIN AND HISTORY OF THE PROBLEM

Three broad categories of works are related to this dissertation. Related considerations include power distribution network analysis and design, thermo-mechanical stress in TSV-based 3D ICs, and low power 3D IC design methodologies.

#### *2.1 3D Power Distribution Network Analysis and Design*

Many efforts have been made to assess and solve power delivery problems in 3D ICs. A compact modeling of power delivery network for 3D ICs and ideas to suppress power supply noise to an acceptable level were presented in [7]. Three different TSV topologies for 3D P/G network were explored for better power integrity in [8]; (1) a large single TSV aligned to a C4 bump, (2) multiple TSVs around a C4 bump, (3) and evenly distributed TSVs throughout a die. It was shown that 3D die stacking has a higher impact on IR-drop than  $Ldi/dt$  noise [9]. 3D stacking inherently increases the resistance of a 3D P/G network due to P/G TSVs which directly impact IR-drop. On the other hand,  $Ldi/dt$  noise due to time varying activities in the modules is mainly determined by off-chip inductive components. A simultaneous TSV optimization scheme for both power and thermal integrity was proposed in [10], where a minimum TSV placement density in unit design area is determined by power and thermal noise sensitivity analysis. However, all these modeling and optimization techniques are based on the regular TSV placement. Additionally, no work considered TSV RC variation impact on the power supply noise.

#### *2.2 Thermo-Mechanical Stress in TSV-based 3D ICs*

Due to the coefficients of thermal expansion (CTE) mismatch between TSV fill material such as copper and silicon substrate, thermo-mechanical stress is induced during fabrication process. This thermo-mechanical stress can affect device performance and cause mechanical failures in the 3D interconnect. The impact of TSV-induced stress on mobility variations

in electron and hole as well as full-chip timing was studied in [11]. Regarding mechanical reliability, the crack growth behavior caused by TSV-induced stress was presented in [12–14]. However, most previous works focused on modeling the thermo-mechanical stress and reliability of a single TSV in isolation.

Moreover, most previous works on the thermo-mechanical stress and reliability of TSV-based 3D ICs have been done separately in chip or package domain. M. Nakamoto *et al.* showed the significant impact of package components on the chip domain stress. They proposed a stress exchange file to transfer the boundary conditions from package-level to silicon-level analysis [15]. However, all of these approaches require FEA methods which are computationally expensive or infeasible for full-chip or -package analysis.

### ***2.3 Low Power 3D IC Designs***

With 3D integration technology, both the average and maximum distance between components can be substantially reduced by placing them on different dies, which translates into significant savings in delay and power. However, all these savings cannot be achieved without careful managements of TSVs, circuit partitioning, and floorplan. Especially, 3D IC is mainly targeting mobile computing applications such as smart phones which need lower power consumption and high data bandwidth.

Bryan Black *et al.* demonstrated two different approaches for implementing high-performance 3D processors [16]. The first approach is stacking memory on logic (Memory+Logic), and the second is implementing a microarchitecture across two or more dies (Logic+Logic). For an Intel Pentium 4-based microprocessor, with Logic+Logic stacking, about 25% of pipeline stages were eliminated, leading to about 15% performance improvement. In addition, fewer buffers, a smaller clock grid, and significantly less global wire yields a 15% power reduction.

Yuh-Fang Tsai *et al.* explored different 3D design options of partitioning a cache [17]. This paper examined possible partitioning approaches for caches designed using 3D structures and presented a delay and energy model to explore different options of partitioning

a cache across different device layers. Because of the size of TSVs, SRAM cell level partitioning is not feasible. Thus, their focus is on sub-array-level partitioning, namely 3D divided wordline (3-DWL) approach and 3D divided bit line approach (3-DBL). For four active device layers, the energy savings of 31.38% is claimed for a 4MB cache with 25nm technology.

## CHAPTER III

### POWER-DELIVERY NETWORK ANALYSIS AND OPTIMIZATION FOR 3D ICS

Power delivery is believed to be one of the biggest challenges in 3D ICs. With the rapid advance of fabrication technology and the increase in number of gates in a unit chip area, the power consumption of a chip has increased dramatically. As multiple dies are stacked together into a smaller footprint, delivering current to all parts of the 3D stack while meeting the noise constraints becomes challenging. This is mainly because the number of TSVs available for power/ground (P/G) nets is limited [18]. In addition, the impact of TSV RC variation on power-supply noise in 3D-power-delivery network (3D PDN) has not been studied well.

In this chapter, a non-regular P/G TSV placement algorithm is presented to reduce the number of P/G TSVs while achieving the given IR-drop noise requirement. Next, TSV RC variation caused by process variation is modeled, and this TSV RC variation impact on both static noise (IR-drop) and dynamic noise (voltage droop) in 3D IC layouts is examined.

#### **3.1 Backgrounds**

##### **3.1.1 Existing Works**

In general, the objective of P/G TSV optimization is to minimize power noise with minimum number of P/G TSVs. Previous works on 3D power delivery networks employed regular P/G TSV placement or optimized the density of P/G TSVs in each P/G tile to meet power noise requirement.

A physical model of 3D power distribution network is presented in [7]. Their model assumed that power is fed from the package through power I/O bumps distributed over the bottom-most die and travels to the upper dies using TSVs and solders. Therefore, P/G TSV locations are predetermined by regularly placed power I/O bumps. Three different TSV topologies for 3D P/G network have been explored in terms of power integrity in [8];

(1) a large single TSV aligned to a C4 bump, (2) multiple TSVs around a C4 bump, (3) and evenly distributed TSVs throughout a die. Again, density and location of P/G TSV were predetermined.

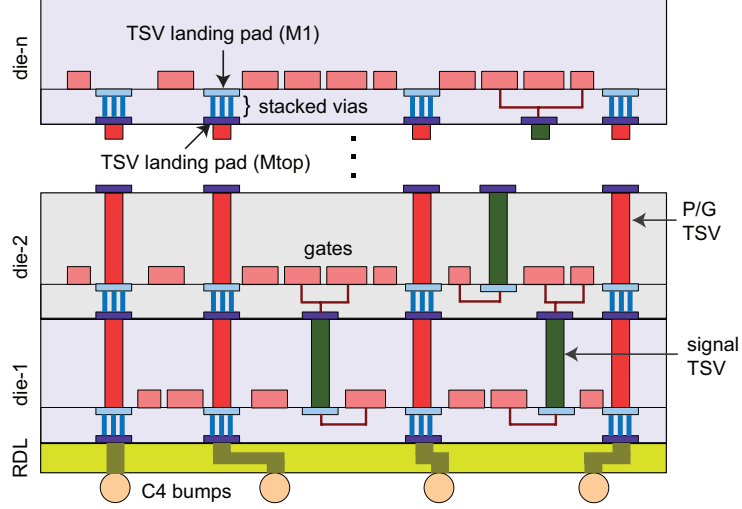
It is shown that 3D die stacking has a higher impact on IR-drop than  $Ldi/dt$  noise [9]. 3D stacking inherently increases the resistance of a 3D P/G network due to P/G TSVs which directly impacts IR-drop. On the other hand,  $Ldi/dt$  noise due to time varying activities in the modules is caused by dominant off-chip inductive components. They also examined the effects of P/G TSV spacing as well as C4 bump spacing on power noise. However, these approaches also assumed regularly placed P/G TSVs with predefined density.

A simultaneous TSV optimization scheme for both power and thermal integrity is proposed in [10]. They first divide each die into  $N$  tiles and define possible TSV placement density. Depending upon the power and thermal noise level in each tile, a minimum TSV density pattern is selected. However, this work only considers noise in P/G planes in the package without considering on-chip power supply routing.

### 3.1.2 P/G TSV Impacts on 3D IC Layouts

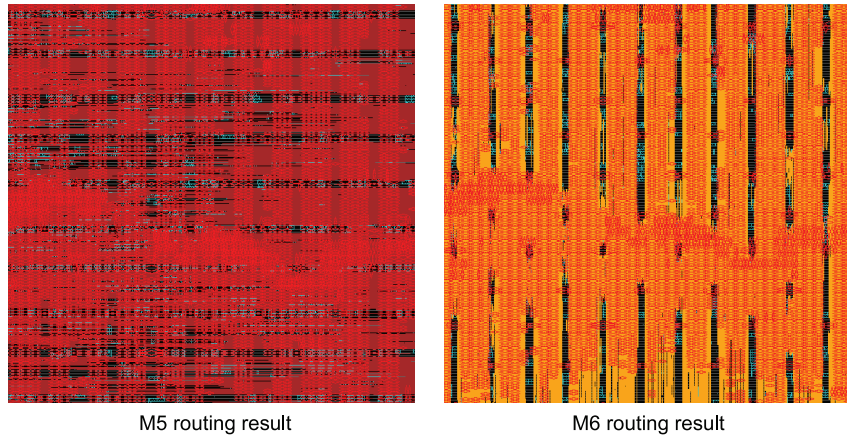
The target 3D structure is illustrated in Figure 1. It is assumed that adjacent dies are bonded in a face-to-back (F2B) fashion. Via-first TSVs interfere with a device layer, whereas via-last TSVs interfere with both device and metal layers. In this study, both P/G TSVs and signal TSVs are the via-first type. Therefore, signal TSVs affect a device layer, the top-most metal layer (M6), and the bottom-most metal layer (M1). However, P/G TSVs are routed through stacked local vias in each die as shown in Figure 1. Thus, these P/G TSVs affect all metal layers and a device layer in a similar way as via-last TSVs. Thus, P/G TSVs cause severe routing congestions if many 3D connections are required. Note that power C4 bumps are connected to P/G TSV landing pads at M6 using redistribution layer (RDL).

A part of signal net routing result of M5 and M6 for a FFT circuit (256 point and 8-bit precision) is shown in Figure 2. P/G TSVs are placed regularly with a  $50\mu m$  pitch on top of P/G bumps in this design. This figure clearly shows that not only the space that P/G TSVs



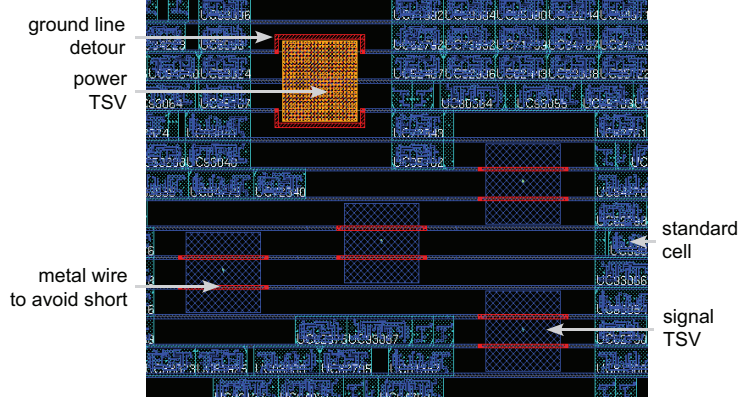
**Figure 1:** Target 3D structure with via-first TSVs. P/G TSVs are vertically connected with stacked local via arrays.

occupy, but space between P/G TSVs are not fully exploited for signal net routing, hence causes more severe routing congestions than expected. In the case of M6 (yellow), wires are routed in vertical direction, and space between P/G TSVs in vertical direction is not used well since horizontal space between P/G TSVs limits the routing capacity for vertical M6 wires. This phenomenon mostly occurs in higher metal layers which are typically used for long signal net connections. It is possible that regularly placed P/G TSVs make a bottleneck for long wires to route. This might cause wirelength increase and performance degradation. If routing space is not enough, foot print area needs to increase to mitigate routing problems.



**Figure 2:** Routing congestion in M5 and M6 due to regularly placed P/G TSVs

A P/G net routing for standard cell rows also becomes challenging. Figure 3 shows a part of P/G net routing result for the same FFT circuit. If the P/G TSV size is larger than a standard cell height, it is inevitable that a single power (ground) TSV also covers the region that ground (power) nets are supposed to be routed. Therefore, power (ground) nets should detour ground (power) TSVs to avoid short between power and ground, which is an additional source of routing congestion. Moreover, P/G nets should also avoid short with an M1 landing pad of a signal TSV. Thus, P/G net routing in 3D ICs consumes more routing resources compared to 2D ICs, hence reduces available routing resources for signal net routing.



**Figure 3:** Power distribution network layout. Additional routing resources are needed for P/G net to detour P/G TSVs and signal TSVs.

These observations call for P/G TSV count reduction. In a chip design phase, power consumption profiles can be estimated based on simulations and power library for standard cells and macro blocks. Thus, it is possible to identify the region that consumes more power and demands more current accordingly than other regions. These high power consuming locations are susceptible to IR-drop noise violation. If IR-drop noise is estimated accurately based on the given power profile, the power noisy spots can be identified. By placing P/G TSVs more in these power noisy spots than other regions, the IR-drop noise can be efficiently reduced while using minimum number of P/G TSVs as well as saving routing resources compared to conventional regularly placed P/G TSVs.

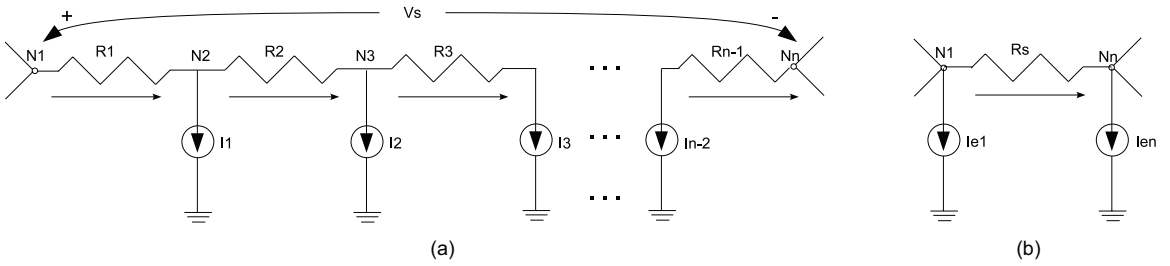


### 3.2 Non-regular Power/Ground TSV Placement Algorithm

P/G rings are routed on the periphery of circuits, and P/G stripes that provide power and ground for each standard cell are routed horizontally. Therefore, a series resistor chain along P/G stripes can be built with current sources that represent standard cells based on a given circuit layout and a power profile. IR-drop estimation is performed on this resistive circuit. To handle large circuits with millions of nodes, an equivalent circuit modeling method is adopted [19, 20]. The proposed P/G TSV placement algorithm is applied to this simplified P/G network to obtain the optimal P/G TSV locations.

#### 3.2.1 Equivalent Circuit for Series Resistors

One example of a resistive P/G network is shown in Figure 4. Consider a series resistor chain in the P/G network in Figure 4(a). There will be some voltage  $V_s$  between the two series ends,  $N_1$  and  $N_n$ . Conceptually, a voltage source with  $V_s$  can be added between the nodes  $N_1$  and  $N_n$  without disturbing the network.



**Figure 4:** Resistive P/G network. (a) Series resistor chain in P/G network. (b) Series resistor equivalent circuit.

Suppose the positive current direction for resistive branch  $R_i$  is from  $N_i$  to  $N_{i+1}$ . A superposition can be applied to this network to produce an equivalent circuit as shown in Figure 4(b), where the positive current direction of  $R_s$  is from  $N_1$  to  $N_n$ . The equivalent resistor  $R_s$  is the sum of all the resistors in series.

$$R_s = \sum_{i=1}^{n-1} R_i. \quad (1)$$

The superposition method is used to determine how the current from each current source divides between the two ends. All current sources except the one in question are replaced

by an open circuit, while the voltage source between nodes  $N_1$  and  $N_n$  is replaced by a short circuit. The resulting system is a current divider, and the additional current at  $N_1$  and  $N_n$  is sum of all the divided currents. The equivalent current  $I_{e_1}$  and  $I_{e_n}$  can be calculated as follows [19, 20]:

$$I_{e_1} = \sum_{i=1}^{n-2} \frac{\sum_{j=i+1}^{n-1} R_j}{R_s} I_i \quad (2)$$

$$I_{e_n} = \sum_{i=1}^{n-2} \frac{\sum_{j=1}^i R_j}{R_s} I_i. \quad (3)$$

Once the network has been solved with the equivalent circuit, the intermediate node voltages and currents are calculated based on the superposition method as follows:

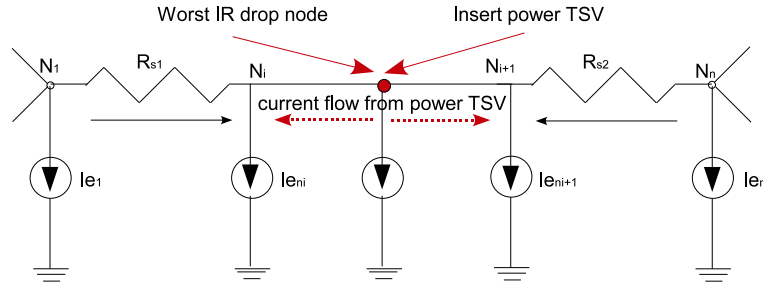
$$V_{i+1} = V_i - \frac{R_i}{R_s} V_s - R_i I_{e_i} \quad (4)$$

$$I_{e_{i+1}} = I_{e_i} - I_i. \quad (5)$$

The equivalent circuit modeling method was originally developed for the fast P/G network simulation [20]. To apply this method for the proposed P/G TSV placement algorithm, following differences need to be considered:

1. In a series resistor chain circuit, not all the node voltage information needs to be considered for possible P/G TSV placement locations. Only the nodes whose voltages are the local minimum (worst IR-drop) or the local maximum (worst ground bounce) need to be considered. Since each power (ground) stripe in a design will have a single local minimum (maximum) node, each P/G stripe will be divided into two sub-chain circuits. Currents will flow in the same direction in each sub-chain, and only the voltages at two ends need to be considered.
2. When P/G TSVs are inserted in the worst IR-drop node, the current flow direction is changed in affected P/G stripes. Even though the total current demands for these P/G nets are unchanged, the current direction is altered with the additional current flow from P/G TSVs. Therefore, the local minimum or the local maximum nodes in these P/G stripes are changed; hence node voltages in these stripes should be updated.

3. Inserting P/G TSVs in some P/G stripes changes both the node voltages at these P/G stripes and the boundary voltages and currents at the ends of these P/G stripes. These changes affect the boundary voltages and currents of other P/G stripes as well. A simple example is shown in Figure 5. Before inserting a power TSV, currents are flowing to the worst IR-drop node. After inserting the power TSV, the amount of current flowing from both ends will be reduced depending on the amount of current that power TSV provides. This will change the amount of current that outer power ring supplies, hence IR-drop through the power ring and the boundary voltages and currents in adjacent power stripes will be changed as well.



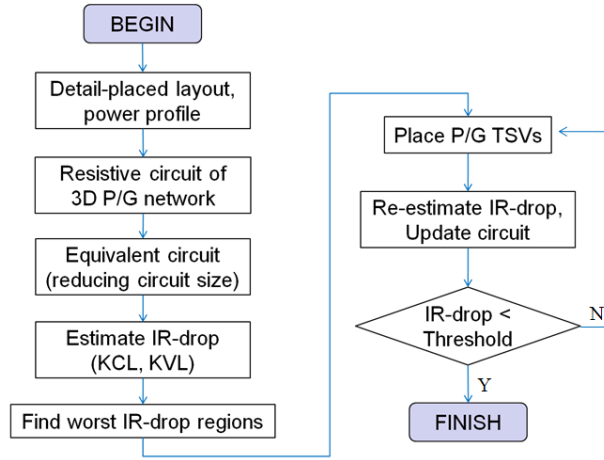
**Figure 5:** Boundary voltage and current change due to P/G TSV insertion.

### 3.2.2 Non-regular P/G TSV Placement Algorithm

The P/G network is constructed based on the detailed cell placement results along with the power profile for each stacked die. Then, the IR-drop noise is calculated based on the Kirchhoff's voltage law (KVL) and current law (KCL). Then, the local minimum and the local maximum node voltages for power and ground net are identified. With this information, the simplified circuit is built to handle large size circuits efficiently using the equivalent circuit model. Next, P/G TSVs are inserted where IR-drop constraint is violated, and IR-drop is re-evaluated using the equivalent circuit model. If the current design meets the target IR-drop threshold, the P/G TSV placement algorithm finishes. If not, additional P/G TSVs are inserted to the current IR-drop violating regions.

To validate the IR-drop estimation algorithm, the results for both 2D and 3D IC layouts

are compared with the results using existing 2D commercial-grade tools, Cadence VoltageStorm [21]. Using the IR-drop estimation method, both 2D and the two die-stacked 3D IR-drop results match the VoltageStorm results within 7% error. The computed resistance value based on the interconnect technology file for each P/G wire segment overestimated by 6% compared with VoltageStorm. Since this is a deterministic error, resistivity values are tuned to match the results. Because of the tool’s limitation on the number of layers it can process ( $= 15$  metal layers), the proposed algorithm is validated up to two die-stacked 3D ICs.



**Figure 6:** Flow chart of the non-regular P/G TSV placement algorithm.

### 3.3 Experimental Results of Non-Regular P/G TSV Placement

The proposed non-regular P/G TSV placement algorithm has been implemented in C language. Four FFT circuits are used for analyses, which are listed in Table 1. All circuits are synthesized using Synopsis Design Compiler [22] with the physical cell library for the target 130nm technology, and designed using Cadence SoC Encounter [23] to 2D and two die-stacked 3D ICs.

The number of signal TSVs is chosen to cover around 10% of the chip area, and the overall placement density including both standard cells and TSVs is targeted to 80%. The TSV diameter, height, and resistance are  $6\mu m$ ,  $30\mu m$ , and  $30m\Omega$ , respectively, which are similar to the data of manufactured TSVs in [24]. Note that the TSV cell size is large, which occupies five standard cell rows (one standard cell row height is  $3.69\mu m$ ).

**Table 1:** Benchmark Circuits.

Circuit	# gates	# signal TSVs	TSV coverage (%)	Clock frequency (MHz)	Profile
FFT1	200K	794	8.9	200	256 point 8 bit precision
FFT2	405K	1397	8.2	142	256 point 16 bit precision
FFT3	910K	7089	16.3	111	512 point 16 bit precision

The IR-drop analysis results between regularly placed P/G TSVs and the proposed algorithm are compared using two-die-stacked 3D designs. The 3D design whose P/G TSVs are regularly placed with a  $100\mu\text{m}$  pitch is used as a baseline as shown in Table 2. The IR-drop constraint for the proposed algorithm is set to be the worst IR-drop noise of the baseline. With the P/G TSV locations obtained from the proposed algorithm, two-die-stacked 3D ICs are designed. Then, the IR-drop noise is validated by Cadence VoltageStorm.

**Table 2:** Results of the regular P/G TSV placement.

Ckt	# P/G TSV (core/peri)	Area ( $\mu\text{m} \times \mu\text{m}$ )	WL (mm)	IR-drop (mV)
FFT1	481 / 170	$1776 \times 1776$	16056	71
FFT2	1013 / 230	$2444 \times 2444$	36240	153
FFT3	2665 / 370	$3902 \times 3902$	87289	236

It is observed that similar IR-drop numbers are achieved using the proposed algorithm with a much smaller number of P/G TSVs compared with the baseline 3D designs as shown in Table 3. The number of P/G TSVs is saved by 59.3% on average. For instance, the P/G TSV count is reduced by 68.4% in FFT3. In addition, as the circuit size becomes larger, the P/G TSV count reduction increases. It is possible that the regular P/G TSV placement scheme uses more P/G TSVs than necessary, especially in non-power-noisy spots. If an accurate power profile is available in a design phase, the number of TSVs for 3D-power-delivery network can be reduced significantly. With the reduced number of P/G TSVs, footprint area and total wirelength are reduced by 3.4% and 3.5% on average, respectively.

### 3.4 TSV RC Variation

The TSV RC variation impact on the 3D PDN is now investigated. Process variations on TSVs are inevitable because of a misalignment, variations of TSV diameter/height and oxide thickness, and wafer surface cleanliness and roughness. However, the extreme misalignment

**Table 3:** Results of the non-regular P/G TSV placement algorithm. Numbers in parentheses are compared with the regular TSV placement case.

Ckt	# P/G TSV (core/peri)	Area ( $\mu m \times \mu m$ )	WL (mm)	IR-drop (mV)	IR-drop % error	runtime (sec)
FFT1	160/170 (49.4% ↓)	$1744 \times 1744$ (2.6% ↓)	15028 (6.4% ↓)	75	5.6	2.63
FFT2	266/230 (60.0% ↓)	$2402 \times 2402$ (3.5% ↓)	35152 (3.1% ↓)	148	3.3	6.76
FFT3	592/370 (68.4% ↓)	$3821 \times 3821$ (4.2% ↓)	86507 (1.0% ↓)	251	6.4	16.44

that causes a systematic variations and increases contact resistance is highly unlikely in the state-of-the-art wafer-bonding processes [25]. Thus, the TSV RC variation can be modeled as random effects. In this work, the TSV RC variation is modeled based on the TSV dimension variation using analytical models.

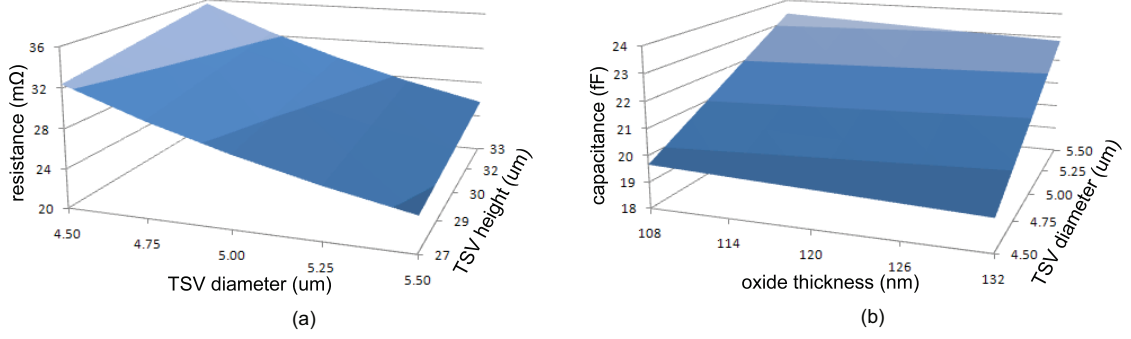
#### 3.4.1 $R_{TSV}$ variation

The analytical expression of the dc resistance of a TSV is given by

$$R_{TSV} = \frac{\rho l_{TSV}}{\pi r_{TSV}^2}, \quad (6)$$

where  $\rho$  is the resistivity of a conducting material, and  $r_{TSV}$  and  $l_{TSV}$  are radius and height of a TSV, respectively. The resistivity of a Cu TSV is  $16.8n\Omega \cdot m$  at  $20^\circ C$ . The contact resistivity of  $0.45\Omega \cdot \mu m^2$  is adopted from the measured data based on the Cu direct bonding [26]. Thus, the total TSV resistance is the sum of the TSV dc resistance and the contact resistance.

The baseline TSV diameter, height, and oxide thickness are  $5\mu m$ ,  $30\mu m$ , and  $120nm$ , respectively. Then, both TSV diameter and height are varied by  $\pm 10\%$  of nominal values to model the process variation. The TSV RC variation range caused by TSV dimension variations is shown in Figure 7. A superlinear relationship between the TSV resistance and the TSV diameter is observed, while a linear dependency is observed between the TSV resistance and the TSV height as shown in Figure 7(a). With the fixed TSV height of  $30\mu m$  and the TSV diameter variation of  $\pm 10\%$  from the nominal value, the TSV resistance changes from -13.6% to +19.3% compared with the nominal TSV resistance.



**Figure 7:** TSV RC variations caused by TSV dimension variations. (a) TSV resistance vs. TSV diameter and height variation. (b) TSV capacitance vs. TSV diameter and oxide thickness variation.

### 3.4.2 $C_{TSV}$ variation

The nature of the TSV  $C$ - $V$  characteristics is similar to the planar MOS capacitor such that the accumulation capacitance is the oxide capacitance given as [27]

$$C_{TSVacc} = C_{ox} = \frac{2\pi\epsilon_{ox}l_{TSV}}{\ln\left(\frac{t_{ox}+r_{TSV}}{r_{TSV}}\right)}. \quad (7)$$

As the TSV bias increases, the depletion capacitance acts in series with the oxide capacitance. This depletion resistance is given by

$$C_{TSVdep} = \frac{2\pi\epsilon_{si}l_{TSV}}{\ln\left(\frac{t_{ox}+r_{TSV}+d_{dep}}{t_{ox}+r_{TSV}}\right)}, \quad (8)$$

where  $t_{ox}$  is the TSV oxide thickness, and  $d_{dep}$  is the depletion width in the silicon substrate. The substrate doping is assumed to be  $2 \times 10^{15}/cm^3$ .

The effective TSV capacitance is the series combination of the oxide capacitance and the depletion capacitance. This effective TSV capacitance is given by

$$C_{TSV} = \frac{C_{ox}C_{TSVdep}}{C_{ox} + C_{TSVdep}}. \quad (9)$$

Although the TSV height has a direct impact on the capacitance value, the deviation of the TSV height from its nominal value is smaller compared with the TSV diameter and the oxide thickness. This is mainly because the TSV height is several times larger than the TSV diameter. Thus, the impact of variations in the TSV diameter and the oxide thickness on the TSV capacitance is examined in this study. Both the TSV diameter and the oxide

thickness are varied by  $\pm 10\%$  from their nominal values. It is observed that the impact of the oxide thickness variation on the TSV capacitance is negligible as shown in Figure 7(b). With the fixed oxide thickness and the TSV diameter variation of  $\pm 10\%$  from its nominal value, the TSV capacitance changes from  $-8.6\%$  to  $+8.6\%$  compared with the original TSV capacitance.

### ***3.5 3D Power Supply Noise Analysis with TSV RC Variation***

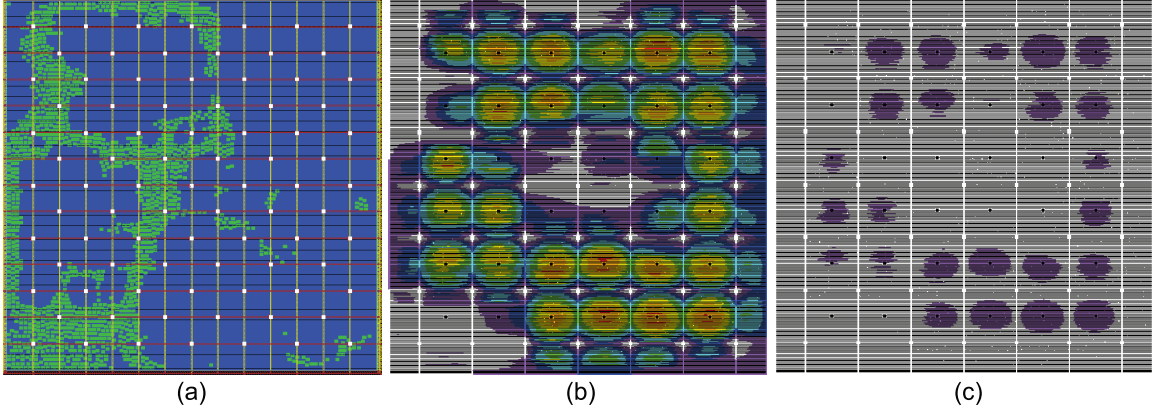
In this study, a sign-off level 3D power supply noise analysis flow is built using existing 2D commercial tools and in-house scripts. First, RC parasitics of P/G net (SPEF) of each die are extracted using Synopsys StarRC [28]. In addition, the coupling capacitance between top metal wires and the substrate of an adjacent die is modeled by performing capacitance extraction on 3D structures using ANSYS Q3D extractor [29]. Then, capacitance values are scaled accordingly for these top metal P/G wires in SPEF file which is not supported by 2D parasitic extraction tools. Then, an in-house tool is employed to merge P/G nets from multiple dies to build a single PDN and P/G TSV parasitics which follow given normal distribution are inserted between adjacent dies.

In addition, design and interconnect technology files of an individual die are modified to create a unified 3D design. Next, pseudorandom input vectors which obey a given switching activity are generated. With 3D design files and input vectors, current wave forms of all signal nets in the 3D design are obtained using Synopsys NanoSim [30]. Then, Synopsys PrimeRail [31] is employed to perform both static and dynamic noise analysis of 3D PDN. Due to the limitation on the number of layers these tools can process, simulations are restricted to two-die stacked 3D ICs. Figure 8 shows the layout and power supply noise maps from the proposed analysis flow.

### ***3.6 Experimental Results with TSV RC Variations***

Three industrial circuits are used for the analysis. All circuits are designed with a NanGate 45nm cell library [32] to two-die stacked 3D ICs, which are listed in Table 4. Top two metal layers are used to construct regularly distributed power grid for each 2D tier and P/G TSVs are inserted at each grid node as shown in Figure 8(a). Decaps are deployed





**Figure 8:** Layout and power supply noise maps in the bottom die of ind1. (a) Layout. White rectangles are P/G TSVs and green rectangles are signal TSVs. (b) Dynamic noise map. Red color represents noisy spots. (c) Static noise map.

uniformly across the die with  $0.5fF/\mu m^2$  density and the worst dynamic noise is reduced by  $53mV$  for the circuit ind1 in a deterministic simulation, for example.

It is assumed that P/G TSVs are connected directly to C4 bumps. The power to power pitch of TSV, C4 bump, and grid is  $100\mu m$  in all benchmark designs. To model C4 bump and package impedance,  $5m\Omega$  resistance,  $500pH$  inductance, and a parallel  $30fF$  capacitance for each C4 bump are used. Current wave forms are generated using 500MHz clock frequency with a switching activity of 0.2 for input signals. The TSV with  $5\mu m$  diameter,  $50\mu m$  height, and  $120nm$  thick oxide is employed. Finally, more than 400 Monte Carlo power supply noise simulations are performed for every case.

**Table 4:** Benchmark circuits for TSV RC variation impact study.

circuit	# gates	area ( $\mu m^2$ )	# signal TSVs	# P/G TSVs	% area by P/G TSVs	% area by decap
ind1	355K	$810^2$	1632	85	0.91	5.32
ind2	1.16M	$1850^2$	14957	613	1.26	5.32
ind3	2.52M	$2822^2$	22413	1405	1.24	5.32

### 3.6.1 Impact of TSV RC Variation Range

First, the impact of TSV RC variation range on the power supply noise in 3D PDN is investigated. TSV RC parasitics which follow the normal distribution with standard deviation of 10, 20, or 30% of a nominal value are generated. As shown in Table 5, negligible variations

from mean values are observed for both static and dynamic noise cases with given RC variation ranges. This is because TSV RC values are much smaller compared with parasitics from 2D P/G grid, decaps, and C4 bumps, hence TSV RC variation does not affect the quality of 3D PDN.

**Table 5:** Impact of TSV RC variation range.

TSV capacitance variation (unit: $fF$ )	TSV resistance variation (unit: $m\Omega$ )							
	$N(45.6, (0.1 \times 45.6)^2)$				$N(45.6, (0.3 \times 45.6)^2)$			
	worst static noise ( $mV$ )		worst dynamic noise ( $mV$ )		worst static noise ( $mV$ )		worst dynamic noise ( $mV$ )	
	mean	% $\sigma/m$	mean	% $\sigma/m$	mean	% $\sigma/m$	mean	% $\sigma/m$
$N(35.5, (0.1 \times 35.5)^2)$	10.562	0	133.055	0.00014	10.562	0	133.055	0.00030
$N(35.5, (0.2 \times 35.5)^2)$	10.562	0	133.055	0	10.562	0	133.055	0.00024
$N(35.5, (0.3 \times 35.5)^2)$	10.562	0	133.055	0.00015	10.562	0	133.055	0.00029

### 3.6.2 Impact of Number of Variation Sources

Since the TSV RC variation range shows negligible effect on power supply noise, the impact of the number of variation sources are explored, i.e. P/G TSV count, to see if increasing number of variation sources worsens the robustness of PDN. The P/G TSV pitch, C4 bump pitch, and P/G grid density are kept same for all circuits to study the impact of P/G TSV count only. Table 6 shows that the number of variation sources poses negligible effects on power supply noise. This is again parasitics of P/G TSVs are much smaller than those of 2D P/G grid.

**Table 6:** Impact of number of variation sources.

circuit	# P/G TSVs	TSV $R \sim N(45.6, (0.3 \times 45.6)^2)$ (unit: $m\Omega$ ) TSV $C \sim N(35.5, (0.3 \times 35.5)^2)$ (unit: $fF$ )			
		worst static noise ( $mV$ )		worst dynamic noise ( $mV$ )	
		mean	% $\sigma/m$	mean	% $\sigma/m$
ind1	85	10.562	0	133.055	0.00029
ind2	613	8.794	0	104.435	0
ind3	1405	9.234	0	121.042	0.00031

### 3.6.3 Impact of Number of C4 Bumps

It is known that P/G C4 bump pitch is a critical factor for 3D PDN quality. In this section, the number of C4 bumps is varied with fixed P/G TSV count and location to explore the

its impact on power supply noise variation. The ind1 circuit is used for this experiment. Both mean and standard deviation of power supply noise increase with decreasing number of bumps as shown in Table 7. This is because P/G TSVs that are directly connected to bumps carry more current than others, hence more susceptible to TSV RC variation. However, power supply noise variation is still negligible with different number of C4 bumps.

**Table 7:** Impact of number of P/G C4 bumps.

# bumps	TSV $R \sim N(45.6, (0.3 \times 45.6)^2)$ (unit: $m\Omega$ ) TSV $C \sim N(35.5, (0.3 \times 35.5)^2)$ (unit: $fF$ )			
	worst static noise ( $mV$ )		worst dynamic noise ( $mV$ )	
	mean	% $\sigma/m$	mean	% $\sigma/m$
21	25.948	0.058	242.897	0.070
42	12.630	0.029	133.864	0.001
85	10.56	0	133.05	0.0002

#### 3.6.4 Impact of TSV Size

So far, the TSV diameter of  $5\mu m$  is used for all experiments. Now the impact of TSV size on power supply noise variation is investigated. The TSV diameter of 2.5, 5, and  $10\mu m$  with an aspect ratio of 10 are used for this. In addition, the circuit ind2 is redesigned with different TSV sizes, since the TSV size can affect layout quality significantly. Table 8 shows that increasing TSV size expands footprint area significantly and accommodates more P/G TSVs if we keep the P/G TSV pitch same. Although mean values of both static and dynamic noise change due to TSV size, variation of power supply noise is still negligible.

It is natural that fabrication technology will scale down the TSV size to increase TSV density. However, smaller TSV size might cause more problems in PDN due to increased TSV resistance and its variation. To evaluate the impact of TSV RC parasitics of future nano-scale TSVs, power supply noise analysis is performed with the circuit ind1. TSV height is kept as  $15\mu m$  for all cases. Table 9 shows that variation of power supply noise is still small compared with mean values, even though the magnitude of standard deviation increases. More serious problem is high noise level itself in nano-scale TSVs, not the variation in power supply noise.

**Table 8:** Impact of TSV size. Aspect ratio is 10 for all TSVs.

TSV diameter ( $\mu m$ )	area ( $\mu m^2$ )	# P/G TSVs	TSV RC $\sim N(m, (0.3m)^2)$			
			worst static noise (mV)		worst dynamic noise (mV)	
			mean	% $\sigma/m$	mean	% $\sigma/m$
2.5	1610 <sup>2</sup>	421	11.762	0.00177	166.166	0
5	1850 <sup>2</sup>	613	8.794	0	104.435	0
10	2430 <sup>2</sup>	1005	10.193	0.00094	203.645	0.00122

**Table 9:** Impact of parasitics of nano-scale TSVs.

TSV diameter ( $\mu m$ )	TSV RC variation range	worst static noise (mV)		worst dynamic noise (mV)	
		mean	% $\sigma/m$	mean	% $\sigma/m$
1	$\sigma = 0.1m$	10.562	0	133.053	0
	$\sigma = 0.2m$	10.562	0	133.053	0
	$\sigma = 0.3m$	10.562	0	133.053	0
0.5	$\sigma = 0.1m$	14.706	0.08822	154.170	0.07671
	$\sigma = 0.2m$	14.705	0.11981	154.163	0.10521
	$\sigma = 0.3m$	14.701	0.17602	154.124	0.15435
0.1	$\sigma = 0.1m$	157.872	0.24234	847.569	0.22812
	$\sigma = 0.2m$	157.877	0.32161	847.942	0.31513
	$\sigma = 0.3m$	157.877	0.43053	847.568	0.42943

### 3.7 Summary

In this chapter, the impacts of P/G TSVs on IR-drop noise as well as 3D IC layouts are explored. Due to the large size of TSVs, both signal and P/G net routing in 3D IC becomes challenging. The non-regular P/G TSV placement algorithm is proposed to minimize the number of P/G TSV used while satisfying IR-drop noise constraint. Experimental results show that the proposed non-regular P/G TSV placement algorithm reduces the number of P/G TSVs by 59.3% as well as footprint area and wirelength compared with the conventional regular P/G TSV placement scheme.

In addition, the effect of TSV RC variation on the robustness of two-die stacked 3D PDN is examined. Simulation results show that TSV RC variations cause negligible influence on both static and dynamic noise in 3D PDN due to much smaller RC parasitic values of TSVs compared with that of entire PDN. This indicates that building robust 3D PDN under nominal condition is important.

## CHAPTER IV

### TSV STRESS-AWARE FULL-CHIP MECHANICAL RELIABILITY ANALYSIS AND OPTIMIZATION FOR 3D ICS

Due to the coefficients of thermal expansion (CTE) mismatch between TSV fill material such as copper and silicon substrate, thermo-mechanical stress is induced during fabrication process. This thermo-mechanical stress can affect device performance [11] or drive crack growth in 3D interconnects [12,13]. Most previous works focused on modeling the thermo-mechanical stress and reliability of a single TSV in isolation. These simulations have been performed using FEA methods which are computationally expensive or infeasible for full-chip analyses.

In this chapter, a full-chip TSV thermo-mechanical stress and reliability analysis flow is presented which overcomes the limitation of the FEA method. In addition, a design optimization methodology is proposed to reduce mechanical reliability problems in TSV based 3D ICs. This tool is then extended to perform a thermo-mechanical stress and reliability analysis under material property variations. This variation impact on electron and hole mobility as well as the full-chip performance are also examined. This comprehensive study encompasses in-depth modeling of individual TSV and device characteristics, efficient method for large-scale full-chip analysis, and design methods for variation tolerance.

#### ***4.1 Detailed Baseline Modeling***

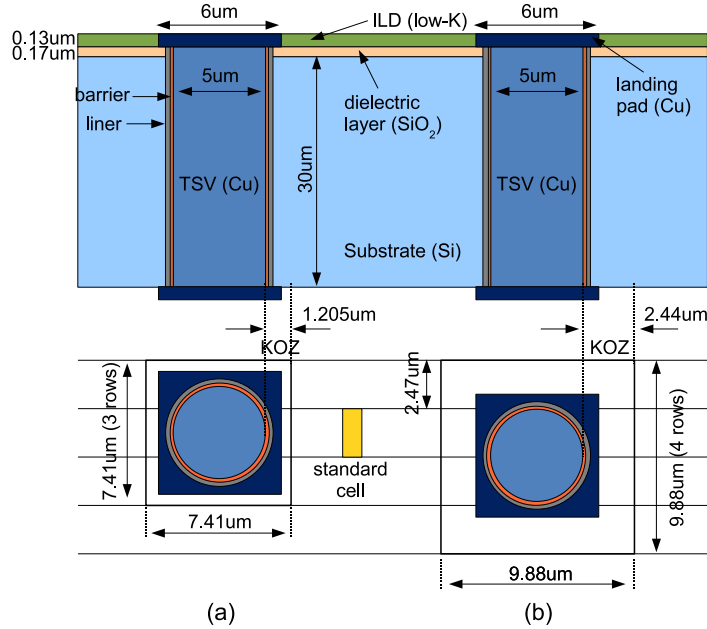
The analytical 2D radial stress model, known as *Lamé* stress solution, was employed to address the TSV thermo-mechanical stress effect on device performance [11]. This 2D solution assumes an infinitely long TSV embedded in an infinite silicon substrate and provides stress distribution in silicon substrate region, which can be expressed as follows [33]:

$$\begin{aligned}\sigma_{rr}^{Si} &= -\sigma_{\theta\theta}^{Si} = -\frac{E\Delta\alpha\Delta T}{2} \left( \frac{D_{TSV}}{2r} \right)^2 \\ \sigma_{zz}^{Si} &= \sigma_{rz}^{Si} = \sigma_{\theta z}^{Si} = \sigma_{r\theta}^{Si} = 0\end{aligned}\tag{10}$$

where  $\sigma^{Si}$  is stress in silicon substrate,  $E$  is Young's modulus (= a measure of stiffness of an elastic material),  $\Delta\alpha$  is mismatch in CTE,  $\Delta T$  is differential thermal load,  $r$  is the distance from TSV center, and  $D_{TSV}$  is TSV diameter.

Even though this closed-form formula is easy to handle, this 2D solution is only applicable to the structure with TSV and substrate only, hence it is inappropriate for the realistic TSV structure with a landing pad and a dielectric liner. It also does not capture the 3D nature of a stress field near the wafer surface around TSVs where devices are located. Moreover, the TSV/substrate interface region near the wafer surface is known to be a highly problematic area for mechanical reliability [13].

Since there is no known analytical stress model for a realistic TSV structure, 3D FEA models for a TSV structure are created to investigate the stress distribution near the wafer surface. To realistically examine the thermo-mechanical stress induced by TSVs, the baseline simulation structure of a TSV is based on the fabricated and the published data [12, 24, 34, 35], as shown in Figure 9.



**Figure 9:** Baseline TSV structure. (a) TSV<sub>A</sub> cell occupying three standard cell rows (KOZ = 1.205μm). (b) TSV<sub>B</sub> cell occupying four standard cell rows (KOZ = 2.44μm).

Two TSV cells are constructed, i.e., TSV<sub>A</sub> and TSV<sub>B</sub>, which occupy three and four standard cell rows in NCSU 45nm technology [36]. The keep-out-zone (KOZ) in which no

cell is allowed to be placed is defined as  $1.205\mu m$  and  $2.44\mu m$  from TSV edge for TSV<sub>A</sub> and TSV<sub>B</sub> cells, respectively. The baseline TSV diameter, height, Cu diffusion barrier thickness, liner thickness, and landing pad size are  $5\mu m$ ,  $30\mu m$ ,  $50nm$ ,  $125nm$ , and  $6\mu m$ , respectively, unless otherwise specified, which are close to the data in [24]. In addition, SiO<sub>2</sub> and Ti are used as a baseline liner and a Cu diffusion barrier material, respectively. Material properties used for our experiments are listed in Table 10. The commercial tool ABAQUS [37] is employed to perform FEA simulations, and all materials are assumed to be linear elastic and isotropic. Also, perfect adhesion is assumed at all material interfaces [38].

**Table 10:** Material properties.

Material	CTE (ppm/K)	Young's modulus (GPa)	Poisson's ratio
Cu	17	110	0.35
Si	2.3	130	0.28
SiO <sub>2</sub>	0.5	71	0.16
Low K	20	9.5	0.3
BCB	40	3	0.34
Ti	8.6	116	0.32
Ta	6.8	186	0.34

#### 4.1.1 3D FEA Simulation

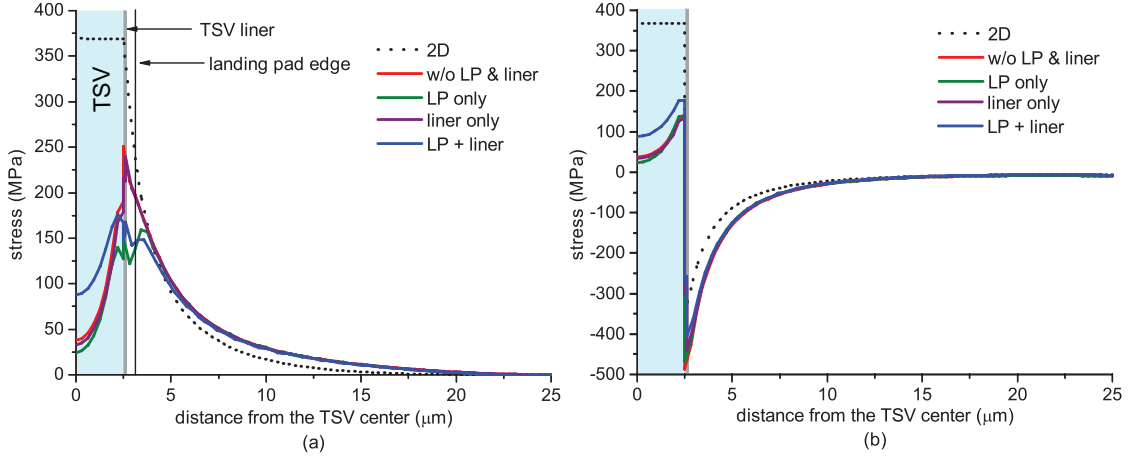
Before discussing the detailed stress modeling results, the concept of a stress tensor is first introduced. Stress at a point in an object can be defined by the nine-component stress tensor:

$$\sigma = \sigma_{ij} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} \\ \sigma_{31} & \sigma_{32} & \sigma_{33} \end{bmatrix}$$

where, the first index  $i$  indicates that the stress acts on a plane normal to the  $i$  axis, and the second index  $j$  denotes the direction in which the stress acts. If index  $i$  and  $j$  are same we call this a normal stress, otherwise a shear stress. Since a cylindrical coordinate system is adopted in this modeling for the cylindrical TSV, index 1, 2, and 3 represent  $r$ ,  $\theta$ , and  $z$ , respectively.

#### 4.1.2 Impact of TSV Liner and Landing Pad

Figure 10 shows FEA simulation results of a normal stress components  $\sigma_{rr}$  and  $\sigma_{\theta\theta}$  along an arbitrary radial line from the TSV center at the wafer surface with  $\Delta T = -250^\circ\text{C}$  of thermal load. That is, we assume the TSV structure is annealed at  $275^\circ\text{C}$  and cooled down to  $25^\circ\text{C}$  to mimic the manufacturing process [13, 33, 39]. It is also assumed that the entire TSV structure is stress free at the annealing temperature.



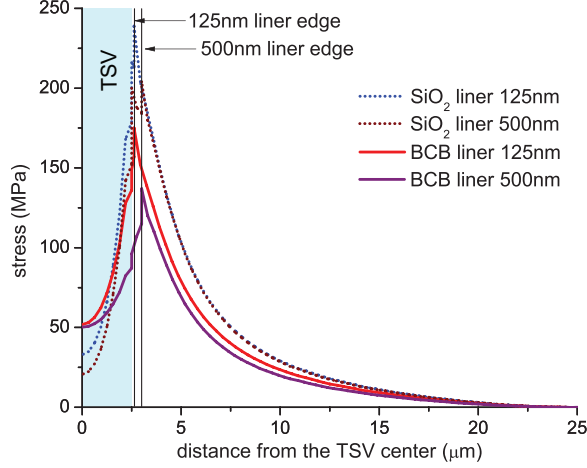
**Figure 10:** Effect of TSV structures on normal stress components. (a)  $\sigma_{rr}$  stress. (b)  $\sigma_{\theta\theta}$  stress.

In 3D FEA simulations, TSV surrounding structures such as a dielectric liner and a landing pad are considered as well, while the 2D model only considers TSV and substrate which are infinitely long in z-direction. Due to this structural difference, the huge discrepancy is observed between 2D solution and 3D stress results at the TSV edge. It is widely known that most of mechanical failures occur at the interface between different materials, hence this TSV edge is the critical region for the reliability. Therefore, the 2D solution does not predict mechanical failure mechanism for TSVs correctly. Also, the  $\text{SiO}_2$  liner which acts as a stress buffer layer reduces  $\sigma_{rr}$  stress at the TSV edge by 35MPa compared with the case without a landing pad and a liner. The landing pad also helps decrease stress magnitude at the TSV edge.

The benzocyclobutene (BCB), a polymer dielectric material, has been employed as an alternative TSV liner material [13, 33]. Since Young's modulus of BCB is much lower



than Cu, Si, and SiO<sub>2</sub>, this BCB liner can absorb the stress effectively induced by the CTE mismatch. The impact of liner material and its thickness on  $\sigma_{rr}$  stress component is shown in Figure 11. As the liner thickness increases, the stress magnitude at the TSV edge decreases noticeably, especially for the BCB liner case.



**Figure 11:** Effect of liner material/thickness on  $\sigma_{rr}$  stress.

It is evident from these simulations that modeling stress distributions considering surrounding structures such as a liner and a landing pad is important to analyze the thermo-mechanical stress around TSVs accurately. A stress library is constructed by varying TSV diameter/height, landing pad size, and liner material/thickness to enable full-chip thermo-mechanical stress and reliability analysis with different TSV structures.

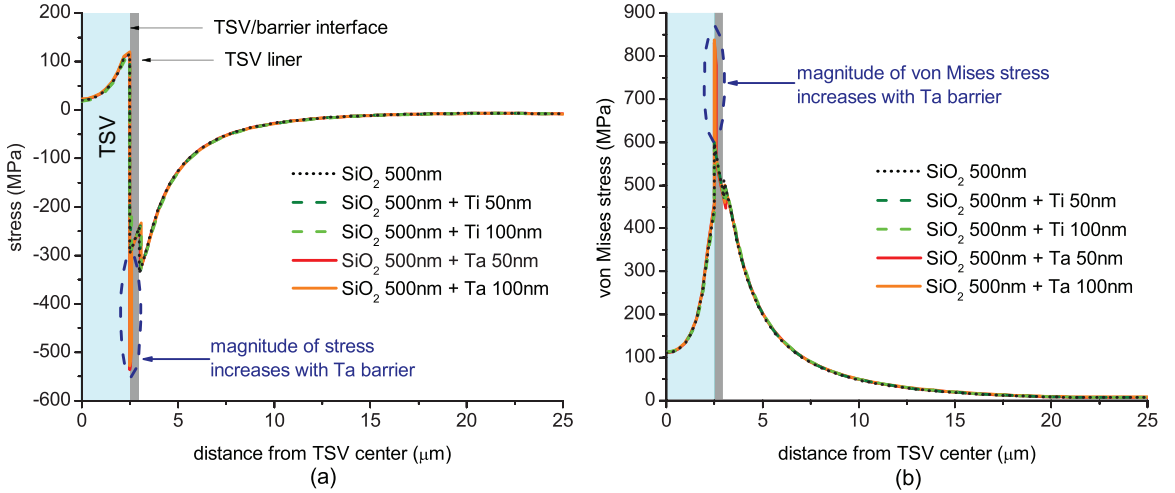
#### 4.1.3 Impact of Cu Diffusion Barrier

For Cu-based interconnects, a barrier layer is needed to prevent Cu diffusion into both dielectrics and Si substrate. This Cu diffusion induces degradation of dielectric layers, hence forms mid-gap defects in Si substrate. These defects serve as a recombination center and reduce the minority carrier lifetime [40]. Therefore, Cu diffusion barrier is deposited between Cu TSV and dielectric liner.

Some previous works ignored Cu diffusion barrier material such as Ti and Ta in FEA simulations, since this barrier thickness is only a small fraction of SiO<sub>2</sub> liner thickness, e.g., one tenth of liner thickness in general. Hence, its impact on stress distribution is negligible [12, 34].

However, thermo-mechanical stress is highly dependent on Young's modulus as well as CTE mismatch. As shown in Table 10, CTE of both Ti and Ta are in between Cu and Si, hence it is unlikely that these barrier materials induce additional stress around TSV on top of stress induced by CTE mismatch between Cu and Si. However, Ta is the stiffest material used in this work. Materials with high Young's modulus cannot absorb stress efficiently. Thus, there is a high chance of stress build-up at this TSV/barrier (Ta) interface.

Figure 12 shows the impact of both Ta and Ti barrier material on the stress and mechanical reliability around TSV. In this experiment, we use a 500nm thick SiO<sub>2</sub> liner, and 50nm and 100nm thick Ta and Ti barrier. We observe a huge increase in stress magnitude in the case of Ta barrier at TSV/barrier interface. For example, in the case of  $\sigma_{\theta\theta}$  stress component shown in Figure 12(a), we see 241MPa and 232MPa increase in compressive stress at TSV/barrier interface for 50nm and 100nm thick Ta barriers respectively, compared with no barrier case. However, there is a negligible change in stress with Ti barrier. This stress increase with Ta barrier also worsens von Mises stress, which is a reliability metric and is discussed in detail in section 4.2.3, as shown in Figure 12(b).



**Figure 12:** Effect of Cu diffusion barrier on stress. (a)  $\sigma_{\theta\theta}$  stress. (b) Von Mises stress.

To further verify that Young's modulus is the key parameter affecting stress magnitude at TSV/barrier interface, Young's modulus of Ta is varied from 25GPa to 225GPa with 25GPa step while CTE and Poisson's ratio are unchanged. As Table 11 shows, as Young's modulus increases, maximum von Mises stress at TSV/barrier interface increases as well.

Therefore, the barrier material for Cu TSV should be chosen carefully to suppress additional mechanical reliability problem on top of existing concerns. Thus, Ti is used as a Cu diffusion barrier material throughout this chapter.

**Table 11:** Effect of Young’s modulus of barrier material on von Mises stress at TSV/barrier interface. Young’s modulus of Ta changes while CTE and Poisson’s ratio remain same.

	Young’s modulus (GPa)								
	25	50	75	100	125	150	175	200	225
stress (MPa)	559	572	575	612	674	740	806	875	943

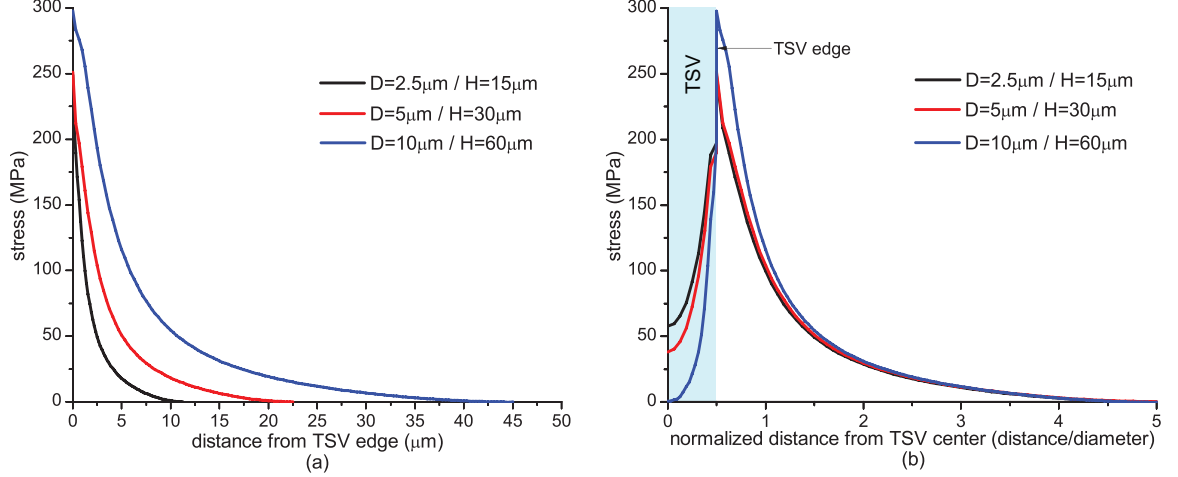
#### 4.1.4 Stress Influence Zone

The magnitude of thermo-mechanical stress induced by TSV is highest at the TSV edge. However, as shown in Figure 10, the magnitude of every normal stress component decays fast, and at around  $25\mu m$  from the TSV center, stress is almost negligible. For an efficient and fast full-chip stress analysis, it is crucial to confine stress analysis to the manageable extent. Thus, a stress influence zone is defined as a circle with a radius of  $25\mu m$  from the TSV center for our baseline TSV with  $5\mu m$  diameter. Beyond the stress influence zone, stress induced by the TSV under consideration is neglected.

The impact of TSV size on stress influence zone is further investigated. Three different TSV diameters with a same aspect ratio of 6 are used; TSV small ( $H/D = 15/2.5 \mu m$ ), TSV medium ( $H/D = 30/5 \mu m$ ), and TSV large ( $H/D = 60/10 \mu m$ ), where  $H/D$  is TSV height/diameter. Figure 13(a) shows the magnitude of  $\sigma_{rr}$  stress component from TSV edge, and the stress magnitude of smaller TSV decays and reaches zero faster.

Figure 13(b) shows stress magnitude along the normalized distance from TSV center, i.e.,  $r/D_{TSV}$ , where  $r$  is the distance from TSV center, and  $D_{TSV}$  is TSV diameter. Even though there are differences in stress magnitude inside TSV and up to  $1 \times$  TSV diameter, stress magnitudes are almost identical beyond that distance. In the 2D solution shown in Equation 10, the magnitude of  $\sigma_{rr}^{Si}$  and  $\sigma_{\theta\theta}^{Si}$  is proportional to  $(D_{TSV}/2r)^2$ . That is why the similar stress curves are observed along the normalized distance.

Most importantly, the stress magnitude becomes negligible at around  $5 \times$  TSV diameter for all three cases. Therefore, the stress influence zone is set as  $5 \times$  TSV diameter in this



**Figure 13:** Effect of TSV size on stress influence zone. (a)  $\sigma_{rr}$  stress from TSV edge. (b)  $\sigma_{rr}$  stress along normalized distance (distance/TSV diameter).

study.

#### 4.1.5 Anisotropic Material Property of Silicon

Up to this point, all materials are assumed to be isotropic for simplicity. However, Si is an anisotropic material with elastic behavior that depends on which crystal direction the structure is being stretched. The possible values of Young's modulus ( $E$ ) for Si range from 130 to 188GPa, and those for Poisson's ratio ( $\nu$ ) range from 0.048 to 0.4. Thus, the choice of this value can affect analysis results significantly [41]. Also, a recent study showed that TSV-induced stress measurement data matches well with FEA simulation results with anisotropic Si material property [42]. In this section, the impact of anisotropic material property of Si on stress distribution is examined compared with the isotropic material property.

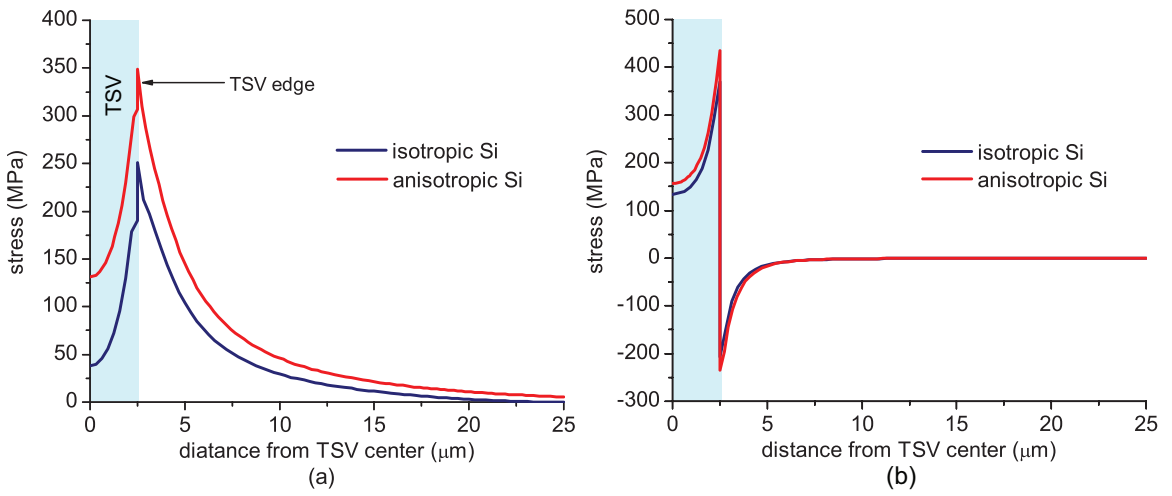
Elasticity is the relationship between stress ( $\sigma$ ) and strain ( $\epsilon$ ). Hooke's law describes this relationship in terms of stiffness  $C$ , i.e.,  $\sigma = C\epsilon$ . For isotropic uniaxial cases, stiffness  $C$  can be represented by a single value of Young's modulus  $E$ . In an anisotropic material, a fourth rank stiffness tensor with  $3^4 = 81$  terms is required to describe the elasticity. Fortunately, due to cubic symmetry of Si, the elastic properties can be expressed in terms of orthotropic material constants. An orthotropic material is one which contains at least two orthogonal planes of symmetry, and Si, with cubic symmetry, can be described this way. The orthotropic elasticity of Si can be expressed with reference axes of a standard

(100) Si wafer, which are  $[110]$ ,  $[\bar{1}10]$ , and  $[001]$ ,

$$\begin{bmatrix} \sigma_{xx} \\ \sigma_{yy} \\ \sigma_{zz} \\ \sigma_{yz} \\ \sigma_{zx} \\ \sigma_{xy} \end{bmatrix} = \begin{bmatrix} c1 & c5 & c6 & 0 & 0 & 0 \\ c5 & c1 & c6 & 0 & 0 & 0 \\ c6 & c6 & c2 & 0 & 0 & 0 \\ 0 & 0 & 0 & c3 & 0 & 0 \\ 0 & 0 & 0 & 0 & c3 & 0 \\ 0 & 0 & 0 & 0 & 0 & c4 \end{bmatrix} \begin{bmatrix} \epsilon_{xx} \\ \epsilon_{yy} \\ \epsilon_{zz} \\ \epsilon_{yz} \\ \epsilon_{zx} \\ \epsilon_{xy} \end{bmatrix}$$

where, orientation specific constants  $c1$ ,  $c2$ ,  $c3$ ,  $c4$ ,  $c5$ ,  $c6$  are 194.5, 165.7, 79.6, 50.9, 35.7, and 64.1, all in GPa, respectively. This stiffness tensor translates to  $E_x = E_y = 169$  GPa,  $E_z = 130$  GPa,  $\nu_{yz} = 0.36$ ,  $\nu_{zx} = 0.28$ , and  $\nu_{xy} = 0.064$  [41].

Figure 14 shows stress comparison between anisotropic and isotropic Si (Young's modulus = 130GPa and Poisson's ratio = 0.28) material properties. The stress magnitude with anisotropic Si is significantly higher than isotropic Si case largely due to higher Young's modulus, especially along X direction as shown in Figure 14(a). It is also observed that the same trend in Y direction ( $= \sigma_{yy}$  stress). However, there is a small difference in Z direction stress component compared with other directions as shown in Figure 14(b). This can be explained by the same Young's modulus  $E_z = 130$  GPa for both anisotropic and isotropic Si.



**Figure 14:** Effect of anisotropic Si material property on stress (a)  $\sigma_{xx}$  stress. (b)  $\sigma_{zz}$  stress.

In general, higher stress level is observed by adopting anisotropic Si material property compared with the isotropic Si case mostly due to higher Young's modulus. Although FEA simulation with isotropic Si provides a good understanding of stress distribution, proper material properties need to be used to accurately assess thermo-mechanical reliability of TSV-based 3D ICs.

## **4.2 *Full-Chip Stress and Reliability Analysis***

FEA simulations of thermo-mechanical stress for multiple TSVs require huge computing resources and time, thus it is not suitable for full-chip analysis. In this section, a full-chip thermo-mechanical stress and reliability analysis flow is presented. First the principle of linear superposition of stress tensors from individual TSVs is examined. Based on the linear superposition method, a full-chip stress map and then von Mises yield metric to predict mechanical reliability problems in TSV-based 3D ICs are built.

### **4.2.1 Linear Superposition Principle**

A useful principle in the analysis of linearly elastic structures is that of superposition. The principle states that if the displacements at all points in an elastic body are proportional to the forces producing them, the body is linearly elastic. The effect, i.e., stresses and displacements, of a number of forces acting simultaneously on such a body is the sum of the effects of the forces applied separately. This principle is applied to compute the stress at a point by adding the individual stress tensors at that point caused by each TSV as follows:

$$S = \sum_{i=1}^n S_i$$

where,  $S$  is the total stress at the point under consideration and  $S_i$  is the individual stress tensor at this point due to the  $i^{th}$  TSV.

### **4.2.2 Stress Analysis with Multiple TSVs**

In this section, the stress analysis flow considering multiple TSVs is described. First, based on the observation that the stress field of a single TSV in isolation is radially symmetric due to the cylindrical shape of a TSV, the stress distribution (= stress tensor) around a TSV

is obtained along an arbitrary radial line from the TSV center in a cylindrical coordinate system. To evaluate a stress tensor at a point affected by multiple TSVs, a conversion of a stress tensor to a Cartesian coordinate system is required. This is due to the fact that stress tensors are extracted from a TSV whose center is the origin in the cylindrical coordinate system; hence a vector sum of stress tensors at a point from each TSV which has a different center location cannot be performed. That is why a universal coordinate system is needed, i.e. Cartesian coordinate system in this case.

Then, a stress tensor at the point of interest is computed by adding up stress tensors from TSVs affecting this point. A TSV stress influence zone is defined as  $25\mu m$  from the center of a TSV with  $5\mu m$  diameter, since the magnitude of stress components becomes negligible beyond this distance, which has been verified by FEA simulations.

Let the stress tensor in Cartesian and cylindrical coordinate system be  $S_{xyz}$  and  $S_{r\theta z}$ , respectively.

$$S_{xyz} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix}, S_{r\theta z} = \begin{bmatrix} \sigma_{rr} & \sigma_{r\theta} & \sigma_{rz} \\ \sigma_{\theta r} & \sigma_{\theta\theta} & \sigma_{\theta z} \\ \sigma_{zr} & \sigma_{z\theta} & \sigma_{zz} \end{bmatrix}$$

The transform matrix  $Q$  is the form:

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where,  $\theta$  is the angle between the X axis and a line from the TSV center to the simulation point. A stress tensor in a cylindrical coordinate system can be converted to a Cartesian coordinate system using conversion matrices:  $S_{xyz} = QS_{r\theta z}Q^T$ .

#### 4.2.3 Mechanical Reliability Analysis

To evaluate if computed stresses indicate possible reliability concerns, a critical value for a potential mechanical failure must be chosen. The von Mises yield criterion is known to be one of the most widely used mechanical reliability metric [43–45]. If the von Mises stress exceeds a yielding strength, material yielding starts. Prior to the yielding strength, the

material will deform elastically and will return to its original shape when the applied stress is removed. However, if the von Mises stress exceeds the yield point, some fraction of the deformation will be permanent and non-reversible.

There is a large variation of yield strength of Cu in the literature, from 225MPa to 600MPa, and it has been reported to depend upon thickness, grain size, and temperature [43]. We use 600MPa as a Cu yielding strength in our experiments. The yield strength of silicon is 7000MPa, which will not be reliability concerns for the von Mises yield criterion. The von Mises stress is a scalar value that can be computed using components of a stress tensor. By evaluating von Mises stress at the interface between TSV and dielectric liner, where highest von Mises stress occurs, we can predict mechanical failures in TSVs.

#### 4.2.4 Validation of Linear Superposition Method

Before using the proposed method, the linear superposition of stress tensors against FEA simulations is validated by varying the number of TSVs and their arrangement. The minimum TSV pitch is set as  $10\mu m$  for all test cases. Stress tensors along a radial line from the TSV center in a single TSV structure (stress tensor list) are obtained through FEA simulation with  $0.1\mu m$  interval. In the linear superposition method, simulation area is divided into uniform array style grid with  $0.05\mu m$  pitch. If the stress tensor at a grid point under consideration is not obtainable directly from the stress tensor list, the stress tensor at the point is computed using linear interpolation with adjacent stress tensors in the list.

Some of comparisons are shown in Table 12. First, a huge run time reduction in the linear superposition method is observed. Note that FEA simulations are performed using 4 CPUs while only one CPU is used for the linear superposition method. Even though the linear superposition method performs stress analysis on a 2D plane at the wafer surface, whereas FEA simulation is performed on entire 3D structure, stress analysis for other planes can be performed in a similar way if needed. Also, the run time in the linear superposition method shows linear dependency on the number of simulation points, which is closely related to the number of TSVs under consideration. Thus, the linear superposition method is highly scalable, hence applicable to full-chip scale stress simulations. More details on scalability is



discussed in section 4.2.7.

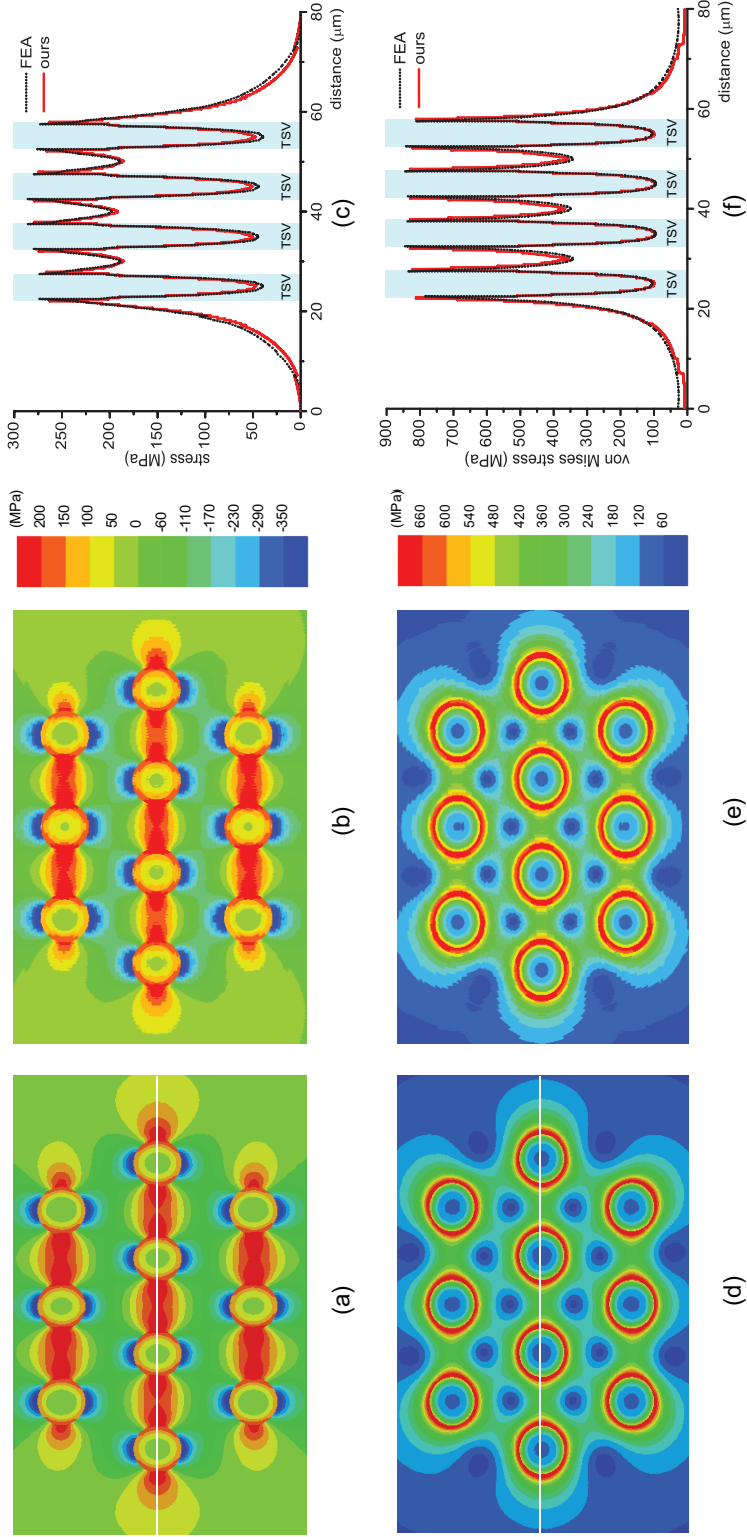
**Table 12:** Von Mises stress comparison between FEA simulations and linear superposition method.

# TSV	FEA		linear superposition		max % error	
	# node	run time	# sim' point	run time	inside TSV	outside TSV
1	153K	21m35s	1.M	20.63s	1.0	-0.4
2	282K	58m11s	1.2M	26.21s	3.3	-0.8
3	358K	1h28m24s	1.44M	36.43s	4.8	-1.3
5	546K	1h59m05s	1.68M	56.02s	12.7	-1.9
10	1124K	4h34m14s	2.24M	65.32s	13.6	-2.0

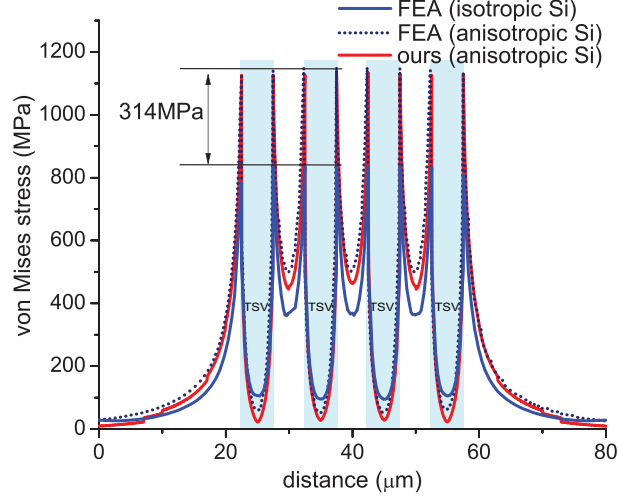
Most importantly, error between FEA simulations and the linear superposition method is practically negligible. Results show that the linear superposition method overestimates stress magnitude inside TSV. However, though maximum % error inside TSV of 10 TSVs case is as high as 13.6%, stress magnitude difference between FEA and the linear superposition method is only 5.0MPa. In addition, since most mechanical problem occurs at the interface between different materials, this error inside TSV does not pose a serious impact on our reliability analysis. Figure 15 shows the stress map of  $\sigma_{xx}$  component and the von Mises stress map for one of test cases which contains 10 TSVs, and it clearly shows the linear superposition method matches well with the FEA simulation result.

Moreover, the feasibility of linear superposition with anisotropic Si material properties is examined. Even though elastic properties of Si depend on the orientation of the structure, fortunately, Young's modulus along X ( $[100]$ ) and Y ( $[\bar{1}10]$ ) direction are same, i.e.,  $E_x = E_y = 169\text{GPa}$ . Since the focus here is the stress on the device layer, which is on the XY-plane, it is possible that the linear superposition method still holds with anisotropic Si elastic properties, though  $E_z$  is different from  $E_x$  and  $E_y$ .

As shown in Figure 16, the linear superposition method still matches well with the FEA simulation result with the anisotropic Si material property. Though, the error increased compared with the isotropic Si case as listed in Table 13, it is still promising to use the linear superposition method for anisotropic Si case. A huge increase in maximum von Mises stress is observed at TSV edge, 314MPa in this case, with anisotropic Si case due to higher Young's modulus in the XY-plane. This is expected as discussed in section 4.1.5.



**Figure 15:** Sample stress comparison between FEA simulation and linear superposition method. (a) FEA result ( $\sigma_{xx}$ ). (b) Ours ( $\sigma_{xx}$ ). (c) FEA vs. ours ( $\sigma_{xx}$ ) along the white line in (a). (d) FEA result (von Mises stress). (e) Ours (von Mises stress). (f) FEA vs. ours (von Mises stress) along the white line in (d).



**Figure 16:** Von Mises stress comparison between isotropic and anisotropic Si. Stress is taken along the white line in Figure 15(d).

**Table 13:** Error in Maximum von Mises stress in MPa (and % error) at TSV edge between FEA simulation and linear superposition for both isotropic and anisotropic Si.

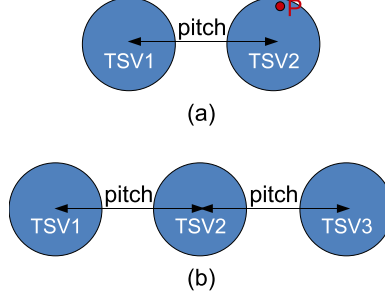
Si property	# TSV (pitch = 10 $\mu m$ )			
	2	3	5	10
isotropic	4.1 (0.5%)	-6.9 (-0.8%)	-5.7 (-0.7%)	-7.0 (-0.9%)
anisotropic	-12.1 (-1%)	-16.3 (-1.4%)	-16.8 (-1.5%)	-18.3 (-1.6%)

#### 4.2.5 Limit of Linear Superposition Method

It is observed that higher % error occurs inside TSV in Section 4.2.4. In this section, the limit of linear superposition method is investigated. In the proposed approach, stress tensors are obtained along a radial line from TSV center to substrate in a single TSV structure through FEA simulation. That is, stress tensors are obtained across at least two different materials. Then, the linear superposition principle is used to compute stress tensor at a point with stress tensors prepared before. This point under consideration could be either inside TSV (Cu) or outside TSV (silicon substrate).

In a full chip, most of simulation points are outside TSV. The stress tensor in the substrate region is computed by adding up stress tensors which are from the substrate region in a single TSV structure. In other words, all the stress tensors in the substrate region are computed by using stress tensors from the substrate region, which is the same material.

However, computing stress tensors inside TSV is different. For example, assume that there is one TSV (TSV1) that affects a point  $P$  inside another TSV (TSV2) shown in Figure 17(a). The stress tensor from TSV1 affects  $P$  is originated from the silicon substrate region, while that from TSV2 affects  $P$  is obtained from Cu. Thus, the stress tensor at  $P$  is calculated by adding stress tensors from two different materials.



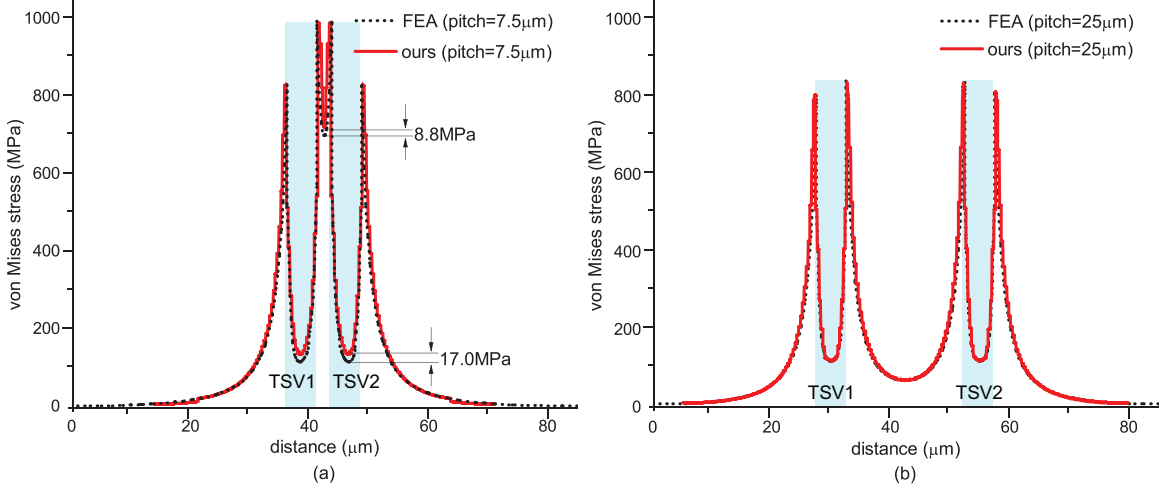
**Figure 17:** Simulation structures for limit of linear superposition method. (a) Two TSVs. (b) Three TSVs.

Even though the linear elastic model is used for the whole structure, due to the difference in material properties between Cu and silicon, stress effect induced by nearby TSVs acting on the TSV under consideration is different between inside and outside the TSV. Therefore, the error from linear superposition is inevitable, especially inside TSV.

Intuitively, if the TSV pitch becomes smaller, the error between FEA simulation and the linear superposition method, i.e., stress (ours) - stress (FEA), will be larger due to high stress magnitude. To further explore this, two TSVs structure is compared by varying TSV pitch from  $7.5\mu m$  to  $25\mu m$  shown in Figure 17(a). Table 14 shows the maximum error between FEA simulation and our method for both inside and outside TSV. For example, at  $7.5\mu m$  pitch, error is 17.0MPa inside TSV and 8.8MPa outside TSV as shown in Figure 18(a). However, this maximum error inside TSV occurs in the TSV center where stress magnitude is lowest, hence the mechanical reliability at this location is not a concern. In addition, the highest error outside TSV is found at the center location between TSVs. Nonetheless, the trend of stress magnitude remains same and the error is not significant. Moreover,  $7.5\mu m$  pitch with  $5\mu m$  diameter TSV is too extreme case for current TSV fabrication process. Therefore, the error induced by the linear superposition method is negligible practically.

**Table 14:** Maximum error (in MPa) in von Mises stress between FEA simulation and linear superposition.

# TSV	location	TSV pitch ( $\mu m$ )				
		7.5	10	15	20	25
2	inside	16.5	7.8	3.3	1.8	1.7
	outside	8.9	2.8	0.7	-1.1	-0.7
3	inside	15.4	6.9	5.5	2.5	1.6
	outside	7.2	2.3	-1.2	-0.8	-0.5

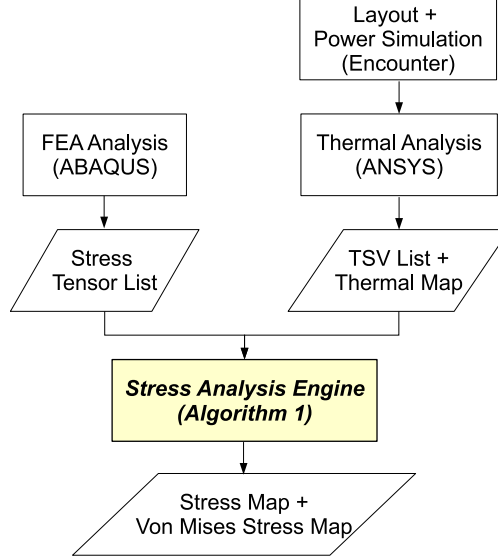


**Figure 18:** Von Mises stress along the center line in Figure 17(a) with two TSVs with different pitches. (a) Two TSVs with pitch of  $7.5\mu m$ . (b) Two TSVs with pitch of  $25\mu m$ .

It is further examined whether an additional TSV worsens error. Three TSVs structure is used shown in Figure 17(b). As similar to two TSVs case, the maximum error occurs at TSV2 center. However, as shown in Table 14, an additional TSV does not aggravate simulation errors. Thus, the linear superposition approach is suitable for full-chip analysis with an acceptable error.

#### 4.2.6 Full-Chip Analysis Flow

The full-chip thermo-mechanical stress and reliability analysis flow is shown in figure 19. First, a detailed FEA simulation of a single TSV is performed to obtain the stress tensors along a radial line from the TSV center. This stress tensor list is an input to the simulation engine. In addition, the locations of the TSVs from a 3D IC layout along with a thermal map are provided to the tool.



**Figure 19:** Simulation flow

The basic algorithm for generating stress and reliability maps is illustrated in algorithm 1. With aforementioned inputs, a stress influence zone from each TSV is determined. Then, the points in the influence zone are associated with the affecting TSV. Next, for each simulation point under consideration, the stress tensor from the TSV found in the association step are looked up, and the coordinate conversion matrices are used to obtain stress tensors in the Cartesian coordinate system. Each TSV affecting this simulation point is visited and their stress contributions are added up. Once the stress computation at a point is finished, the von Mises stress value is calculated. The complexity of this algorithm is  $O(n)$ , where  $n$  is number of simulation points.

#### 4.2.7 Scalability of Algorithm

To verify the scalability of the linear superposition method, the algorithm is applied to 10, 100, 1000, 10000 and 100000 TSVs that are regularly placed across a chip with  $15\mu m$  pitch. Note that the entire chip area is divided into uniform array style grid with  $0.2\mu m$  pitch in these experiments.

Table 15 shows that run time increases almost linearly as the number of simulation points increases. The algorithm visits each simulation point and calculates stress tensor at this point by linearly superposing stress tensors from TSVs located within a stress influence

---

**Algorithm 1:** Full Chip Stress and Reliability Analysis Flow.

---

```
input : TSV list  $T$ , stress library, thermal map (optional)
output: stress map, von Mises stress map
1 for each TSV  $t$  in  $T$  do
2    $c \leftarrow$  center of  $t$ 
3    $r \leftarrow$  FindStressInfluenceZone( $c$ )
4   for each point  $r'$  in  $r$  do
5      $r'.TSV \leftarrow t$ 
6   end
7 end
8 for each simulation point  $p$  do
9   if  $p.TSV \neq \emptyset$  then
10    for each  $t \in p.TSV$  do
11       $d \leftarrow$  distance( $t, p$ )
12       $S_{cyl} \leftarrow$  FindStressTensor( $d, temperature$ )
13       $\theta \leftarrow$  FindAngle( $line\ tp, x\ axis$ )
14       $Q \leftarrow$  SetConversionMatrix( $\theta$ )
15       $S_{Cart} \leftarrow QS_{cyl}Q^T$ 
16       $p.S_{Cart} \leftarrow p.S_{Cart} + S_{Cart}$ 
17    end
18  end
19   $vonMises(p) \leftarrow$  ComputeVonMises( $p.S_{Cart}$ )
20 end
```

---

zone as shown in algorithm 1. Since the number of TSVs affecting a single point is a small constant in most cases, the number of simulation points dominates the entire run time.

It is also observed that simulation time per each grid point increases as the number of TSVs increases, and saturates beyond 10000 TSVs case. At first, the number of TSVs which affect a point under consideration increases as total number of TSVs increases, hence a single point requires more time to compute stress tensor. However, after certain point, increasing total number of TSVs does not affect the simulation time per point. This is due to a finite stress influence zone, and hence a finite number of TSVs affecting a simulation point.

**Table 15:** Scalability to full-chip analysis.

# TSV	area ( $\mu m \times \mu m$ )	# sim' point	run time	time/point
10	$60 \times 60$	90K	3.6s	$40\mu s$
100	$165 \times 165$	680K	25.4s	$37\mu s$
1000	$495 \times 495$	6.1M	5m55s	$58\mu s$
10000	$1515 \times 1515$	57.4M	1h3m	$66\mu s$
100000	$4770 \times 4770$	568.8M	10h12m	$65\mu s$

Note that these test cases in which TSVs are regularly placed across a chip are worst cases in terms of run time. This is because almost every simulation point needs stress computation, since it is highly likely to be within a stress influence zone of a certain TSV. However, in case TSVs are only placed in some part of the chip or irregularly placed, run time can reduce significantly since there could be nonnegligible number of simulation points that are outside of the stress influence zone of any TSVs, which are filtered out in the first *for loop* of our algorithm 1.

### 4.3 Full-Chip Simulation Results

A TSV-aware full-chip stress and reliability analysis flow is implemented in C++/STL. Four variations of an industrial circuit with changes in TSV placement style and TSV cell size are used for analyses, which are listed in Table 16. The number of TSVs and gates are 1472 and 370K, respectively, for all cases. These circuits are synthesized using Synopsys Design Compiler with the physical library of 45nm technology [36], and layouts are obtained using Cadence Encounter. All circuits are designed to two-die stacked 3D ICs.

**Table 16:** Benchmark circuits.

circuit	TSV placement	TSV cell size ( $\mu m \times \mu m$ )	wirelength (mm)	area ( $\mu m \times \mu m$ )
Irreg <sub>A</sub>	Irregular	$7.41 \times 7.41$	9060	$960 \times 960$
Reg <sub>A</sub>	Regular	$7.41 \times 7.41$	9547	$960 \times 960$
Irreg <sub>B</sub>	Irregular	$9.88 \times 9.88$	8884	$1000 \times 1000$
Reg <sub>B</sub>	Regular	$9.88 \times 9.88$	9648	$1000 \times 1000$

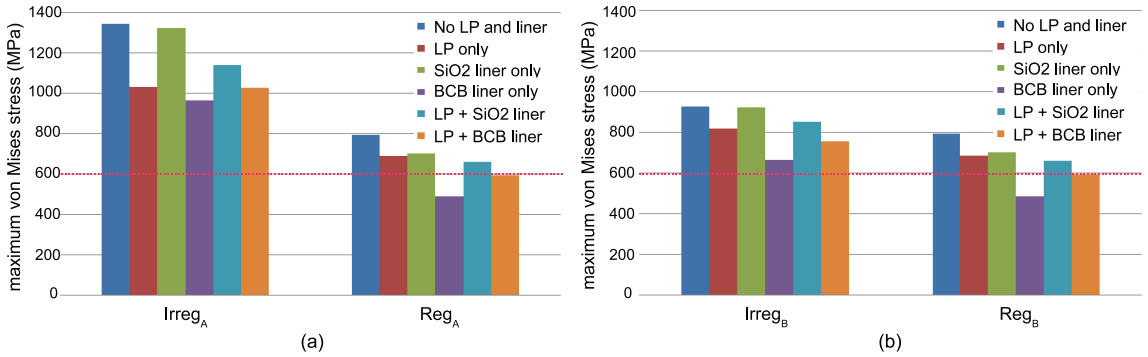
The in-house 3D placer is employed for TSV and cell placement, and details of TSV and cell placement algorithms can be found in [46]. In the regular TSV placement scheme, TSVs are pre-placed uniformly on each die, and then cells are placed, while TSVs and cells are placed simultaneously in the irregular TSV placement scheme. The irregular TSV placement shows better wirelength than the regular case [46]. A gate-level 3D IC design methodology for these circuits is used as a baseline, and these design results are compared with block-level designs in section 4.3.5.



### 4.3.1 Overall Comparison

In this section, the impact of TSV structure, TSV placement style, and KOZ size on the mechanical reliability in 3D ICs are discussed. Full-chip stress and reliability analysis on our benchmark circuits are performed based on our stress modeling results with different TSV structures.

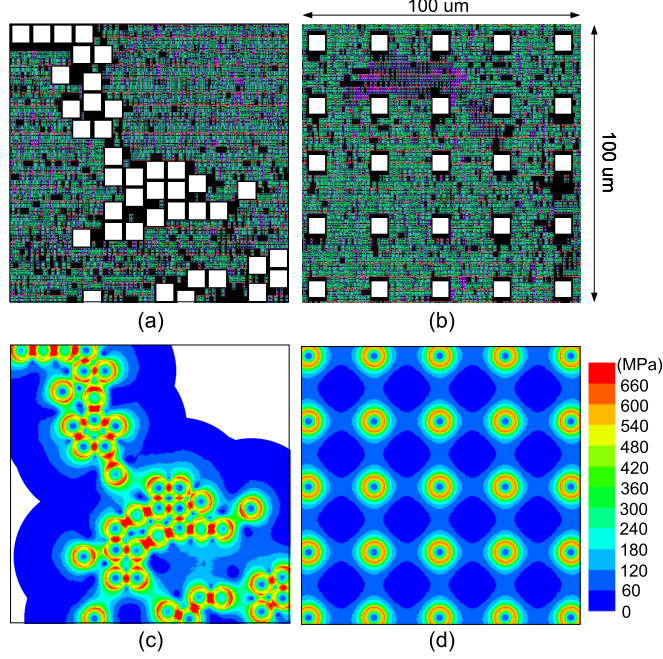
The maximum von Mises stress in our benchmark circuits are shown in Figure 20. First, it is observed that designs with the irregular TSV placement show worse maximum von Mises stress than those with the regular TSV placement. This is mainly because TSVs can be placed closely in case of the irregular TSV placement scheme to minimize wirelength. A small portion of von Mises stress maps of Irreg<sub>A</sub> and Reg<sub>A</sub> circuits is shown in Figure 21, and most of TSVs in the Irreg<sub>A</sub> circuit exceeds the Cu yielding strength (600MPa).



**Figure 20:** Impact of TSV structure, TSV placement style, and KOZ size on the maximum von Mises stress. (a) Designs with TSV<sub>A</sub> cell (KOZ =  $1.205\mu m$ ). (b) TSV<sub>B</sub> cell (KOZ =  $2.44\mu m$ ).

Second, as the KOZ size becomes larger, stress level reduces significantly for the irregular TSV placement case. By enlarging the KOZ size, i.e., increasing TSV cell size in our design flow, TSV pitch increases accordingly. This in turn reduces stress interference between nearby TSVs, and hence decreases von Mises stress level of TSVs. However, for the regular TSV placement case, since the TSV pitch of Reg<sub>A</sub> ( $23.5\mu m$ ) and Reg<sub>B</sub> ( $25\mu m$ ) is similar and also interference from nearby TSVs is negligible at this distance, there is no noticeable difference in maximum von Mises stress.

Third, these results show the importance of using an accurate TSV stress model to assess



**Figure 21:** Close-up shots of layouts and von Mises stress maps. (a) Irreg<sub>A</sub>. (b) Reg<sub>A</sub>. (c) Von Mises stress map of Irreg<sub>A</sub>. (d) Von Mises stress map of Reg<sub>A</sub>.

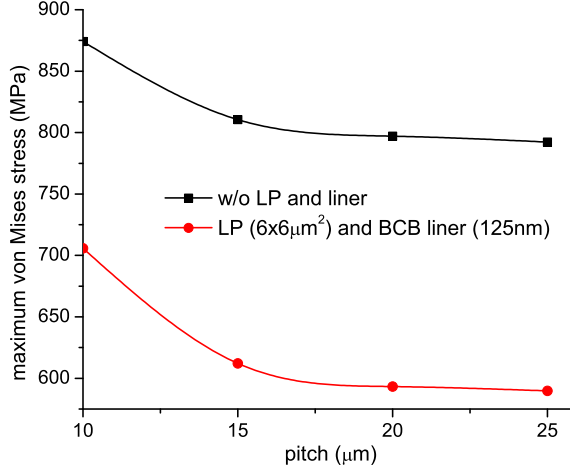
the mechanical reliability of 3D ICs. There are significant differences in the von Mises stress depending on the existence of structures surrounding a TSV, such as a landing pad or a liner. It is possible that the reliability problems might be overestimated by using a simple TSV stress model not considering a landing pad and a liner. However, most of these test cases violate the von Mises yield criterion for Cu TSV.

#### 4.3.2 Impact of TSV pitch

TSV pitch is the key factor that determines stress magnitude in the substrate region between TSVs. In this section, the effect of TSV pitch on von Mises stress is examined. TSVs are regularly placed in  $1 \times 1 \text{ mm}^2$  chip. 1600, 2500, 4356, and 10000 TSVs whose pitch are 25, 20, 15, and  $10 \mu\text{m}$ , respectively, are used. Two data sets are obtained; one without a landing pad, a liner, and a barrier; and another with a  $6 \times 6 \mu\text{m}^2$  landing pad, a  $125 \text{ nm}$  thick BCB liner, and a  $50 \text{ nm}$  thick Ti barrier.

The von Mises stress magnitude decreases with increasing pitch and starts to saturate at around  $15 \mu\text{m}$  pitch as shown in Figure 22. This is understandable since the stress magnitude induced by a single TSV becomes negligible at the similar pitch. In addition, the layout

using TSVs with a landing pad and a BCB liner shows a similar trend with lower von Mises stress magnitude than the case without these structures.



**Figure 22:** Impact of TSV pitch on maximum von Mises stress.

#### 4.3.3 Impact of TSV Size

To investigate the effect of the TSV size, three different sizes of TSV with a same aspect ratio of 6 are employed; TSV small ( $H/D = 15/2.5\mu m$  and KOZ  $1.22\mu m$ ), TSV medium ( $H/D = 30/5\mu m$  and KOZ  $1.202\mu m$ ), and TSV large ( $H/D = 60/10\mu m$  and KOZ  $1.175\mu m$ ), where  $H/D$  is TSV height/diameter. Note that these TSV cells are occupying two, three, and five standard cell rows, respectively, which are selected to minimize the KOZ size difference between them. By setting similar KOZ size, the impact of TSV size can be examined solely. Additionally, the landing pad width is  $1\mu m$  larger than the corresponding TSV diameter, and a  $125nm$  thick  $SiO_2$  liner and a  $50nm$  thick Ti barrier are used for all cases for fair comparisons.

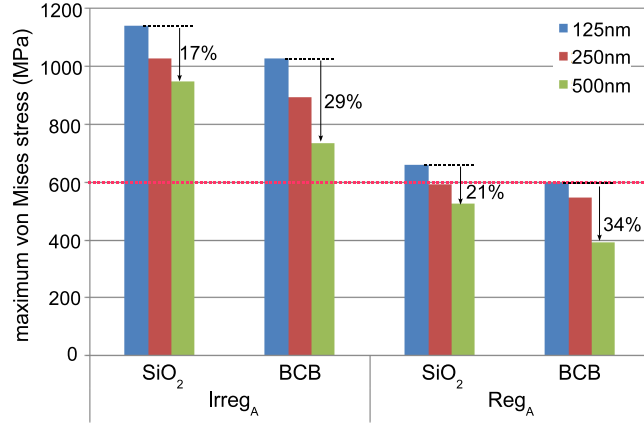
The maximum von Mises stresses are shown in Table 17. Both irregular and regular TSV placement schemes benefit from smaller TSV diameter significantly. This is mainly because the magnitude of normal stress components decay proportional to  $(D/2r)^2$ , where  $r$  is the distance from the TSV center.

**Table 17:** Impact of TSV size on the maximum von Mises stress. The numbers in parentheses are % reduction compared to TSV large case.

TSV placement	max von Mises stress (MPa)		
	TSV large	TSV medium	TSV small
Irregular	1224.6	1126.4 (8% ↓)	902.7 (26% ↓)
Regular	749.3	654.6 (13% ↓)	449.3 (40% ↓)

#### 4.3.4 Impact of Liner Thickness

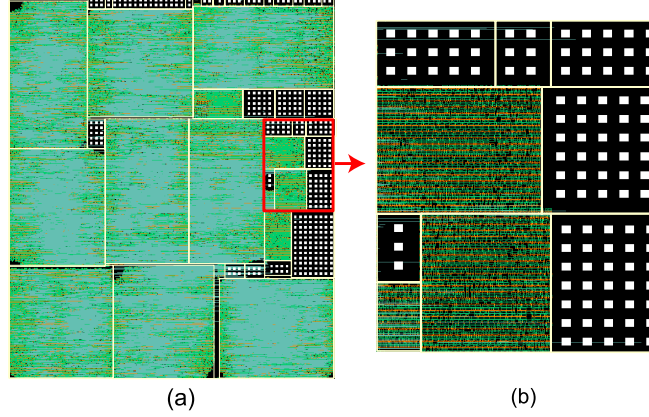
In this section, the impact of liner thickness on von Mises stress is studied. Designs with both TSV<sub>A</sub> cells and TSV<sub>B</sub> cells are employed, and the landing pad size are  $6 \times 6 \mu m^2$  and  $8 \times 8 \mu m^2$ , respectively. In addition, a 50nm thick Ti barrier is used for all cases. Figure 23 shows the maximum von Mises stress results with the liner thickness of 125nm, 250nm, and 500nm.



**Figure 23:** Impact of liner thickness on the maximum von Mises stress of circuits with TSV<sub>A</sub> cell.

It is observed that liner thickness has a huge impact on the von Mises stress magnitude, since the thicker liner effectively absorbs thermo-mechanical stress at the TSV/liner interface. Especially, the BCB liner shows significant reduction in the maximum von Mises stress compared with SiO<sub>2</sub> liner due to extremely low Young's modulus as shown in Table 10. For example, a 500nm thick BCB liner reduces the maximum von Mises stress by 29% for the Irreg<sub>A</sub> and satisfies the von Mises yield criterion for all circuits with a regular TSV placement.

The number of TSVs violating von Mises criterion is listed in Table 18. Even though



**Figure 24:** Layout of block-level design (TSV pitch =  $15\mu m$ ). White rectangles are TSV landing pads. (a) Full-chip layout. (b) Close-up shot of the red box in (a).

there are still many TSVs not satisfying von Mises criterion for the  $Irreg_A$  circuit, it is possible to reduce von Mises stress if TSVs are placed carefully considering this reliability metric during a placement stage.

**Table 18:** Impact of liner thickness on the number of TSVs violating von Mises criterion. The numbers in parentheses are % reduction compared to the  $125\text{ nm}$  thick liner case.

circuit	liner material	# violating TSVs		
		125nm	250nm	500nm
$Irreg_A$	$SiO_2$	1462	1426 (2% ↓)	1281 (12% ↓)
	BCB	1389	1147 (17% ↓)	329 (76% ↓)
$Reg_A$	$SiO_2$	1472	0 (100% ↓)	0 (100% ↓)
	BCB	0	0 (-)	0 (-)
$Irreg_B$	$SiO_2$	1472	1236 (16% ↓)	64 (96% ↓)
	BCB	974	502 (48% ↓)	0 (100% ↓)
$Reg_B$	$SiO_2$	1472	0 (100% ↓)	0 (100% ↓)
	BCB	0	0 (-)	0 (-)

#### 4.3.5 Reliability of Block-Level 3D Design

Even though the gate-level 3D design has the potential of the highest optimization, the block-level design is attractive in the sense that already highly optimized 2D IP blocks can be reused. In this section, the reliability issues in block-level 3D designs are examined. 3D block-level designs are generated using an in-house 3D floorplanner which treats a group of TSVs as a block shown in Figure 24. A  $500\text{ nm}$  thick BCB liner is used for this simulation.

The TSV pitch inside TSV blocks is varied to examine its impact on layout quality as well

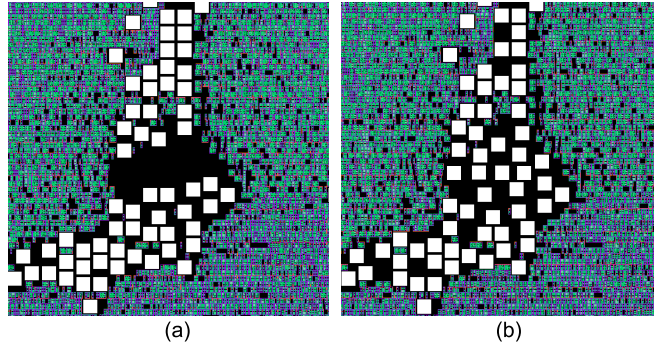
as reliability issues. Table 19 shows that block-level designs use less number of TSVs, show shorter wirelength, and occupy more area than gate-level designs. Experimental results show that the von Mises stress in block-level designs can be controlled with area overhead, since the TSV pitch in block-level design is controllable. Another benefit of block-level design is that the thermo-mechanical reliability problems can be confined to nearby TSV blocks only.

**Table 19:** Comparison between gate-level and block-level design.

design level	TSV pitch ( $\mu m$ )	# TSV	wirelength ( $mm$ )	area ( $\mu m \times \mu m$ )	max stress (MPa)
Gate	irregular	1472	9060	$960 \times 960$	729.3
	23.5	1472	9547	$960 \times 960$	386.4
Block	15	368	8259	$950 \times 1130$	414.5
	10	394	8028	$1080 \times 1000$	519.9
	7.5	333	7933	$980 \times 1090$	716.2

#### 4.3.6 Impact of TSV Re-placement

In this section, TSV locations are manually optimized to show the potential benefit of TSV reliability aware layout optimization while minimizing the change in layout. The Irreg<sub>A</sub> circuit which shows the worst von Mises stress is employed, and a 500nm thick BCB liner is used for this experiment. The related study with this BCB liner on the maximum von Mises stress vs. TSV-to-TSV pitch shows that 10 $\mu m$  pitch is a reasonable choice to reduce von Mises stress considering some safety margin. Thus, densely placed TSVs are repositioned to nearby white spaces if available to reduce the von Mises stress as shown in Figure 25.



**Figure 25:** TSV re-placement to reduce von Mises stress. TSV landing pads are white rectangles. (a) Original layout. (b) After TSV re-placement.

Table 20 shows the distribution of von Mises stress higher than 480 MPa across the die, wirelength, and longest path delay before and after the TSV re-placement. 3D static timing analysis is performed to analyze timing using Synopsys PrimeTime with TSV parasitic information included. It is shown that this TSV re-placement clearly reduces high von Mises stress region. With small perturbations of TSV locations, the number of violating TSVs decreases from 329 to 261, which is 21% improvement with only 0.23% wirelength and 0.81% longest path delay increase, respectively. This small test case shows the possibility of a layout optimization without degrading the original performance too much.

**Table 20:** Impact of TSV re-placement on von Mises stress distribution, wirelength, and longest path delay.

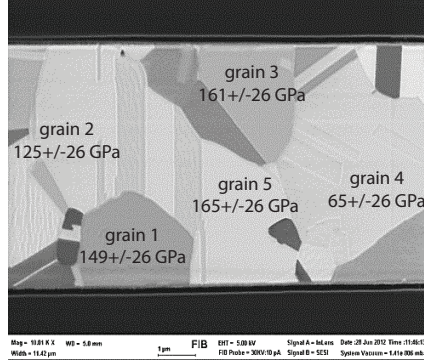
	von Mises stress (MPa)				WL	LPD
	480-540	540-600	600-660	660-	( <i>mm</i> )	( <i>ns</i> )
orig	0.100%	0.041%	0.011%	0.002%	9060	3.607
re-place	0.092%	0.036%	0.009%	0.0%	9081	3.636

#### 4.4 TSV and Device-level Variation Study

Most previous works on the thermo-mechanical stress and reliability of TSV-based 3D ICs assumed a constant material property, i.e., single value for the coefficient of thermal expansion (CTE) and another single value for Young’s modulus of each material in the TSV structure. Thus, each TSV in the 3D IC generates the identical stress distribution around the TSV, and hence the full-chip stress map becomes deterministic [2, 47]. Similarly, the impact of TSV-induced stress on TSV interfacial crack [48] and full-chip 3D timing [11] are limited to this deterministic scenario.

Recently, authors in [4] showed measurement data on the copper (Cu) Young’s modulus variation from 65GPa to 165GPa near the top of a Cu TSV as shown in Figure 26 [4]. This is due to the different sizes and orientations of Cu grains, which are highly process-dependent. They also demonstrated the non-negligible impact of this Cu Young’s modulus variation on electron and hole mobility. In addition, it is well known that the CTE of material is temperature and density-dependent. For example, Cu CTE varies from 15.1ppm/K to 24.2ppm/K over the temperature range from 25°C to 250°C [49]. However, there is no

full-chip stress analysis tool currently available that considers material property variations and their impact in the full-chip scale.



**Figure 26:** Significant Cu Young’s modulus variation due to the variations on grain size and orientation in Cu TSV (focused ion beam (FIB) image with gray contrast [4]).

In following sections, the full-chip thermo-mechanical stress and reliability analysis flow is extended to take into account material property variations. The strategy to handle full-chip complexity is the integration of response surface model (RSM)-based stress computation into the stress linear superposition method [2]. This variation impact on electron and hole mobility as well as the full-chip performance is also investigated. This comprehensive study encompasses in-depth modeling of individual TSV and device characteristics, efficient method for large-scale full-chip analysis, and design methods for variation tolerance.

#### 4.4.1 Simulation Setup

The baseline TSV diameter, height, landing pad size, Cu diffusion barrier thickness, and dielectric liner thickness are  $5\mu m$ ,  $30\mu m$ ,  $6\mu m$ ,  $50nm$ , and  $250nm$ , respectively, which are close to [24]. We use Ti and  $SiO_2$  as Cu diffusion barrier and liner materials. The nominal material properties used for our experiments are as follows: Cu = (17/124), tungsten (W) = (4.5/411), Si = (2.3/130),  $SiO_2$  = (0.5/71), and Ti = (8.6/116), where the first and the second values respectively denote CTE (ppm/K) and Young’s modulus (GPa).

A FEA simulation tool ABAQUS [37] is employed. All materials are assumed to be linear elastic and isotropic [2]. The entire structure undergoes  $\Delta T = -250^\circ C$  of thermal load (annealing  $275^\circ C \rightarrow 25^\circ C$ ) for Cu TSV and  $\Delta T = -400^\circ C$  (deposition/annealing  $425^\circ C \rightarrow 25^\circ C$ ) for W TSV [50]. In addition, all materials are assumed to be stress free at the



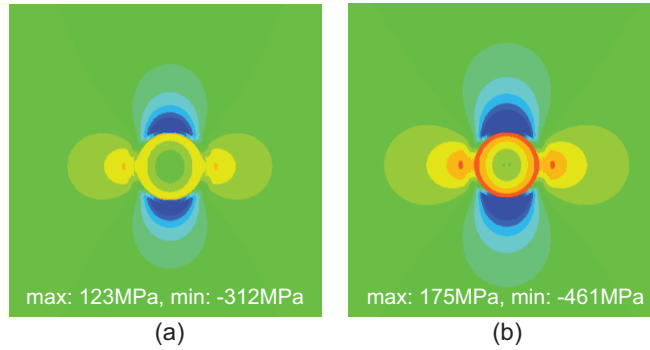
annealing temperature.

#### 4.4.2 Variation Impact on Stress & Reliability

In this section, the impact of each material property on stress is examined and the dominant material properties for stress variations are identified. In the experiments,  $\sigma_{xx}$  and von Mises stress at the device layer are monitored. The significance of these metrics are:

- $\sigma_{xx}$ : normal stress component along the X-direction, which is used as an index for the interfacial delamination between TSV and barrier materials [51].
- The von Mises stress (or equivalent tensile stress): thermo-mechanical reliability metric that is widely used to assess material yielding, i.e., permanent deformation of the structure [2]. In this work, we use this von Mises stress to predict the yielding of TSV material.

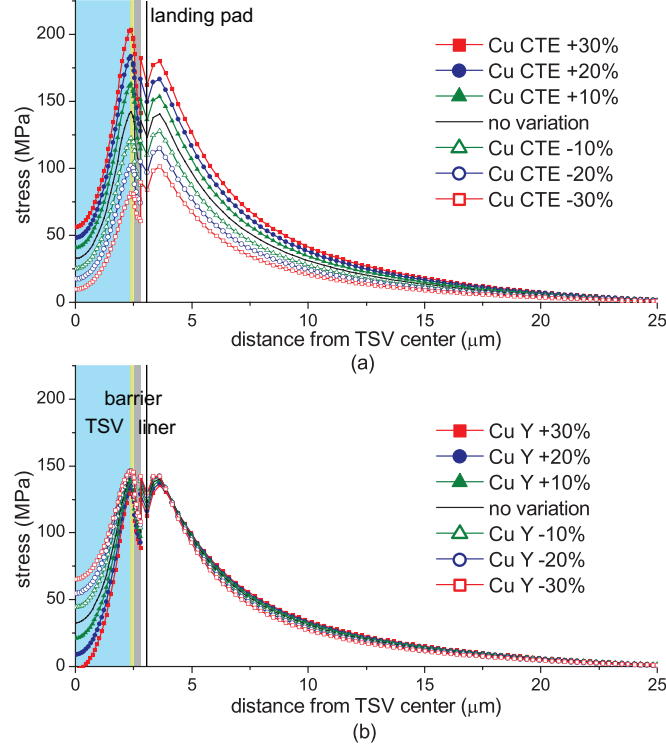
Since TSV-induced stress level is largely determined by the CTE mismatch between TSV material and Si substrate, the CTE of Cu is varied from -15% to +15% from its nominal value. As shown in Figure 27, significant differences in  $\sigma_{xx}$  stress maps between two cases are observed. For example, the maximum tensile stress for -15% and +15% cases are 123MPa and 175MPa, respectively.



**Figure 27:** Cu CTE variation impact on  $\sigma_{xx}$  stress (FEA results). (a) under -15% variation, (b) under +15% variation.

Next, the CTE and Young's modulus of Cu are varied from -30% to +30% from their nominal values and the  $\sigma_{xx}$  stress distributions on the device layer are obtained as shown in Figure 28. A higher Cu CTE causes a larger CTE mismatch between Cu and Si, hence

it elevates the stress level for TSV, barrier, liner, and substrate regions as shown in Figure 28(a). For the extreme +30% case, the maximum  $\sigma_{xx}$  increases by 42% compared with the *no variation* case.



**Figure 28:** Cu variation impact on  $\sigma_{xx}$  under various variation ranges (FEA results). (a) Cu CTE variation, (b) Cu Young's modulus variation.

From Figure 28(b), it is noted that Cu Young's modulus variation also affects the stress distribution for all regions, although its impact is smaller than Cu CTE. The same simulations are also performed for the TSV barrier and liner materials. It is observed that their contributions on the stress distribution are mostly limited within those materials. This is because barrier and liner materials act as a secondary stress contributor in addition to the stress caused by the TSV pillar and substrate. These simulations clearly show that the material property variation in the TSV materials is the major source for the stress variations. However, the material property variations in barrier and liner materials are also modeled to accurately assess the stress inside the barrier and liner as well as the TSV/barrier, barrier/liner, and liner/substrate interfaces.

#### 4.4.3 Variation Impact on Device Mobility

In this section, the impact of material property variation on the hole and electron mobility distributions is investigated. In semiconductors, the changes in inter-atomic spacing resulting from strain significantly affect the band-gaps, making it easier or harder for electrons—depending on the material and strain—to be raised into the conduction band. This results in a change in resistivity of the semiconductor, which also leads to a change in mobility as follows [1]:

$$\begin{aligned} \frac{\Delta R}{R} = -\frac{\Delta\mu}{\mu} = & [\pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy}] \cos^2\phi \\ & + [\pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy}] \sin^2\phi \\ & + \pi_{12}\sigma_{zz} + \pi'_{44}\sigma_{xy}\sin 2\phi \end{aligned} \quad (11)$$

where  $\sigma_{ij}$  is the stress in silicon substrate in Cartesian coordinate system, and  $\phi$  is an angle between the wafer orientation and the transistor channel.

In this work, the (100) Si wafer with reference axes of  $[110]$ ,  $[\bar{1}10]$ , and  $[001]$  is assumed. It is also assumed that the transistor channel direction and the X-axis ( $[110]$ ) are identical. In this setup,  $\pi'_{ij}$  is the piezo-resistivity coefficient defined along the reference axes of (100) Si wafer listed in Table 21:

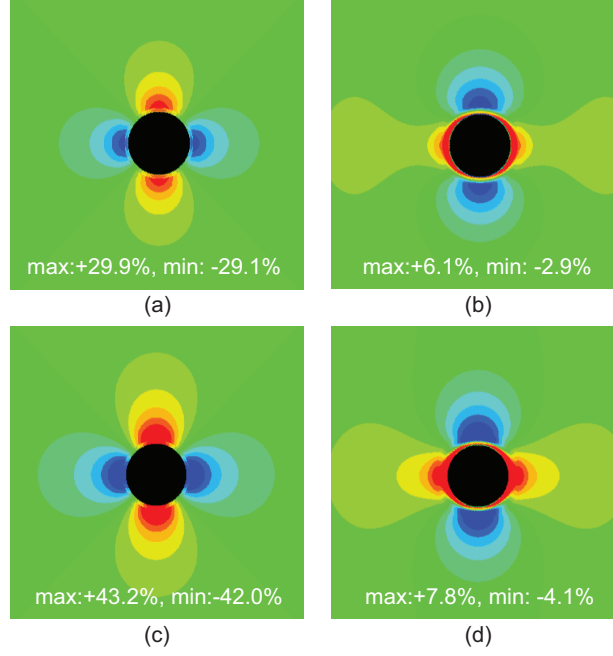
$$\begin{aligned} \pi'_{11} &= \frac{\pi_{11} + \pi_{12} + \pi_{44}}{2} \\ \pi'_{12} &= \frac{\pi_{11} + \pi_{12} - \pi_{44}}{2} \\ \pi'_{44} &= \pi_{11} - \pi_{12} \end{aligned}$$

**Table 21:** Piezo-resistivity coefficient ( $TPa^{-1}$ ) in (100) Si wafer [1].

type	$\pi_{11}$	$\pi_{12}$	$\pi_{44}$	$\pi'_{11}$	$\pi'_{12}$	$\pi'_{44}$
N-type Si	-1022	534	-136	-312	-176	-1556
P-type Si	66	-11	1381	718	-663	77

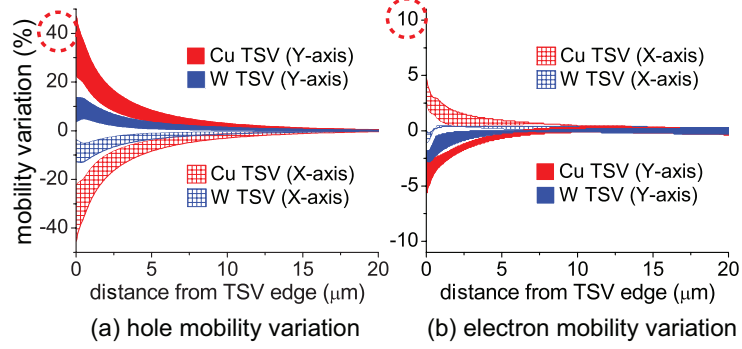
First, Figure 29 shows the mobility variation maps under Cu CTE variations of -15% and +15%. It is shown that higher Cu CTE causes larger mobility variations around the TSV. For example, the maximum hole mobility variation for -15% and +15% cases are

29.9% and 43.2%, respectively. This is a direct consequence of higher stress level as shown in Figure 27. Higher hole mobility variations than electron is also observed. With the same Cu CTE variation range, the minimum and maximum hole mobilities change by 12.9% and 13.3%, respectively, while for the electron mobility, only by 1.2% and 1.7%, respectively.



**Figure 29:** Cu CTE variation impact on carrier mobility. (a) hole mobility map (Cu CTE: -15%). (b) electron mobility map (Cu CTE: -15%). (c) hole mobility map (Cu CTE: +15%). (d) electron mobility map (Cu CTE: +15%).

Next, Figure 30 shows the mobility variation ranges for both Cu and W TSVs under material property variations up to  $\pm 30\%$  of the nominal values. First, the mobility changes significantly depending on the material properties, and the variation range is wider near the TSV. At the Cu TSV edge, the hole mobility changes up to 47%. Thus, the devices near a TSV suffer more performance variations. Also, it is observed that the magnitude and range of mobility variation are much smaller in W TSV because of the smaller CTE of W ( $= 4.5\text{ppm/K}$ ) compared with Cu ( $= 17\text{ppm/K}$ ). For example, at the TSV edge, hole mobility varies between 23% and 47% for Cu TSV, while that changes between 4% and 14% for W TSV. These results indicate non-negligible device performance variations around a TSV under material property variations, which is discussed in detail in Section 4.4.4.



**Figure 30:** Impact of material property variations on mobility. CTE and Young’s modulus of materials in Cu and W TSVs vary from -30% to +30%.

#### 4.4.4 Variation Impact on Cell Delay

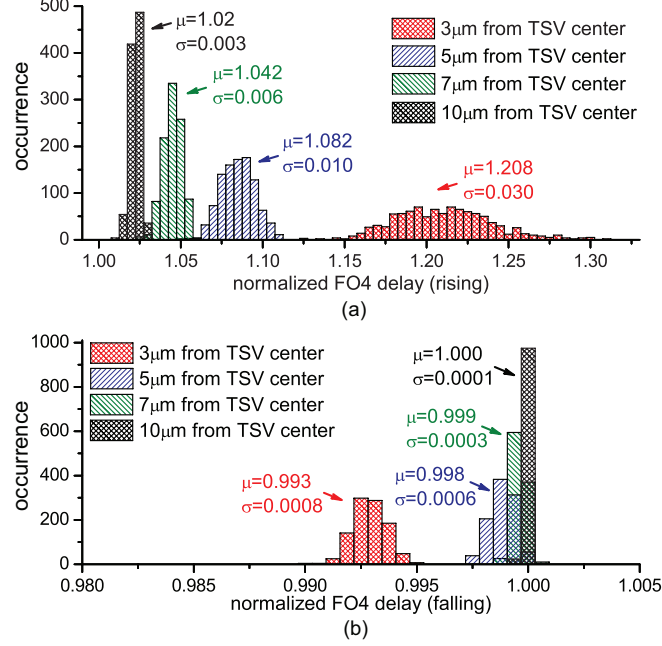
Now the impact of material property variations on the device performance (= cell delay) is studied. A NAND\_X1 gate is placed at four different positions from the TSV center along the X-axis. It is assumed that the transistor channel direction and the X-axis ([110]) are identical in the (100) Si wafer. As Figure 31 shows, cells near the TSV show wider delay distribution as well as higher deviation from the *no variation* case. For example, when the NAND\_X1 gate is placed  $3\mu\text{m}$  away from the TSV center, the mean and maximum FO4 rise delay are 20.8% and 30.7% longer, respectively, than the *no variation* case. In addition, the material property variations affect little on the fall delay as demonstrated earlier in Figure 30(b).

### 4.5 Full Chip-level Variation Study

#### 4.5.1 Challenges

So far, an FEA simulator is used to obtain stress distributions under material property variations. However, in order to conduct *full-chip* stress analysis under variations, either a parametric stress model that considers variations or a full-blown stress library that covers all possible combinations of CTE and Young’s modulus of each material is needed. To build this stress library that covers the variation range up to  $\pm 30\%$  for each material, i.e., varying CTE and Young’s modulus of TSV, barrier, and linear materials from -30% to +30%, even with a 10% step size,  $7^6 = 117649$  FEA simulations are needed, which is impractical.

Another possibility is a linear interpolation of stress tensors between different material



**Figure 31:** Cell FO4 delay variations under material property variations. 1000 Monte Carlo simulations are performed and TSV material is Cu. NAND\_X1 is placed along the X-axis from TSV center. Delay is normalized to *no variation* case. (a) Rise delay distribution. (b) Fall delay distribution.

property variation cases. For example, assume that stress distributions of the Cu CTE +10% and SiO<sub>2</sub> Young's modulus -10% case are obtained by the linear interpolation using the stress tensors from the Cu CTE +10% only case and those from the SiO<sub>2</sub> Young's modulus -10% only case. Then, the number of FEA simulations can be reduced significantly, i.e.,  $6 \times 6 + 1$  (*no variation*) = 37. However, this approach cannot capture the interaction between Cu CTE and SiO<sub>2</sub> Young's modulus variations simultaneously. Mechanical reliability problems occur at the interface between different materials in general, and the stress level at this interface is largely affected by these material properties. Thus, ignoring the interactions between these material properties may not lead to an accurate solution.

#### 4.5.2 Divide-and-Conquer Approach

The Design of Experiment (DOE) method has been widely used in science and engineering applications. It has been proven to be an effective technique when an analysis is desired for complex systems with multiple input factors. It provides a well-organized way of performing experiments so that the experimental results can be used to find meaningful relations

between input factors and responses of the system. In this section, DOE and RSM based stress models that consider material property variations in the TSV structure are presented.

To perform DOE and RSM, design knobs (= input factors) and metrics (= responses) need to be defined first. The stress tensor is our metric. The stress at a point in a body can be described by a nine-component stress tensor, and only six of them are independent. However, only four of these elements are modeled, i.e.,  $\sigma_{xx}$ ,  $\sigma_{yy}$ ,  $\sigma_{zz}$ , and  $\sigma_{xz}$ . This is because the magnitude of  $\sigma_{xy}$  and  $\sigma_{yz}$  is far much lower than others, and hence their impact on the stress distribution is negligible. In the modeling, seven input factors are used; the variation of CTE and Young's modulus of TSV, barrier, and liner materials (= 6), and the distance from the TSV center (= 1).

Note that the stress curve is highly nonlinear as shown in Figure 28. Thus, it is challenging to obtain a smooth response surface that covers the entire region. However, if this is divided into four separate regions, i.e., TSV, barrier, liner, and substrate, RSM-based stress models can be built for each region efficiently. For example, the  $\sigma_{xx}$  stress model for the barrier can be represented as a 1<sup>st</sup> order polynomial as follows:

$$\begin{aligned}\sigma_{xx,barrier} = & c_1 + c_2 \cdot d + c_3 \cdot CTE_{TSV} + c_4 \cdot Y_{TSV} + \\ & c_5 \cdot CTE_{barrier} + c_6 \cdot Y_{barrier} + \\ & c_7 \cdot CTE_{liner} + c_8 \cdot Y_{liner}\end{aligned}$$

where  $d$  is the distance from the TSV center, and  $c_1$  to  $c_8$  are the coefficients for CTE and Young's modulus variation for each material.

60 design points are generated using Stratified Latin Hypercube from space filling design styles. The maximum variation range is set as  $\pm 30\%$  of the nominal value. For each design point, FEA simulation is performed with a single TSV structure and the stress data is obtained. With these stress tensors, the response surface is built and then compact stress models are obtained for four stress tensor components in four separate regions for both Cu and W TSVs.

### 4.5.3 Quality of Stress Models

The goodness-of-fit of a model can be tested with statistics such as coefficient of determination ( $R^2$ ), root mean square error (RMSE), and prediction error sum of squares RMSE (PRESS RMSE), which is evaluated by excluding one data point at a time, building a new RSM model, and computing RMSE [48].

Table 22 shows that the  $R^2$  values of compact stress models are close to 1, and that both RMSE and PRESS RMSE are less than 5MPa for all cases. Considering the fact that stress values near the TSV edge—the critical region for the mechanical reliability—are usually much larger than 100MPa for these four stress components, this fitting quality is acceptable. Even though the models match well with the simulation data, it is essential to validate whether these stress models predict unseen data points correctly. Thus, ten new simulations are designed to validate the stress models. Table 22 shows that the validation RMSE values are close to RMSE values, which confirms that the stress models predict the unseen design points accurately.

### 4.5.4 Compact RSM Model vs. Lookup Table

In Section 4.2.2, stress tensors are extracted along the radial line from the TSV center using FEA simulations and the stress data is stored in a lookup table. Then, the principle of linear superposition is adopted, where the overall stress tensor of a given location is computed by adding the stress tensors from all neighboring TSVs. During this step, each stress tensor is obtained based on the distance between the point and the center of the TSV using the table. However, in this RSM-model-based stress analysis, this table lookup step is eliminated, and hence it saves runtime by around 44% for both a single TSV and 100TSVs cases as shown in Table 23. Figure 32 shows that the RSM-model-based stress analysis result matches well with both FEA and lookup-table-based stress analysis results.

### 4.5.5 Multiple TSVs under Variations

FEA simulations for multiple TSVs require huge computing resources and time, thus it is not feasible for a full-chip-scale analysis. In addition, material property variations for each

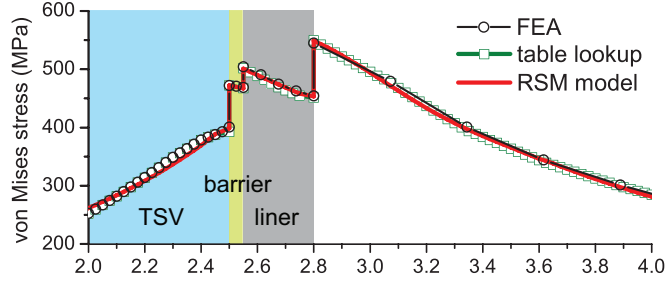


**Table 22:** Quality of the compact stress models for Cu TSV.

	TSV						barrier			liner			substrate					
	$\sigma_z$		$\sigma_x$		$\sigma_y$		$\sigma_z$		$\sigma_x$		$\sigma_y$		$\sigma_z$		$\sigma_x$		$\sigma_y$	
	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{zz}$	$\sigma_{xz}$	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{zz}$	$\sigma_{xz}$	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{zz}$	$\sigma_{xz}$	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{zz}$	$\sigma_{xx}$	$\sigma_{yy}$	$\sigma_{zz}$
$R^2$	0.998	0.996	0.999	0.999	0.999	0.996	0.998	0.997	1.000	0.986	0.993	0.998	0.998	0.992	0.999	0.992	0.998	0.998
RMSE (MPa)	2.105	3.221	3.104	3.894	1.318	3.415	3.345	0.409	2.390	3.938	1.628	1.463	1.486	3.521	4.813	2.332	2.332	2.332
PRESS RMSE (MPa)	2.237	3.341	3.165	3.971	1.412	4.121	4.044	0.509	2.546	4.253	1.804	1.629	1.543	3.624	4.898	2.418	2.418	2.418
Validation RMSE (MPa)	2.221	3.347	3.135	3.871	1.331	3.752	3.574	0.437	2.543	4.014	1.710	1.473	1.502	3.525	4.878	2.348	2.348	2.348

**Table 23:** Comparison between our RSM-model-based analysis and lookup-table-based analysis [2]. TSV pitch is  $10\mu m$ . The numbers in parenthesis are % reduction in runtime compared with the lookup table case.

case		# sims	mesh interval	runtime
1 TSV	FEA	923K	$0.1\mu m$	1h 20m
	lookup table	1.44M	$0.05\mu m$	4.09s
	RSM model	1.44M	$0.05\mu m$	2.26s (-44.7%)
100 TSVs	FEA	-	-	-
	lookup table	4.84M	$0.05\mu m$	290.43
	RSM model	4.84M	$0.05\mu m$	161.35s (-44.4%)



**Figure 32:** Stress comparison between FEA, lookup-table-based analysis [2], and RSM-model-based analysis near TSV edge.

TSV in 3D designs need to be considered. To enable a full-chip stress analysis, the principle of linear superposition of stress tensors from individual TSVs is adopted [2].

The proposed RSM-model-based full-chip thermo-mechanical stress analysis flow is shown in Algorithm 2. This algorithm constructs  $n \times m$  full-chip stress and mobility variation maps from a given full-chip layout and variation conditions. First, each material property for the TSV under consideration is assigned. Since CTE is temperature dependent (linear relationship in general), a full-chip thermal analysis is performed and CTE value for each material is assigned accordingly. Additionally, since this CTE also depends on material density that can be affected by defects such as void, additional variations is imposed (e.g., up to  $\sigma = 10\%$  in our setting) on CTE values. Next, for each point under consideration, the stress tensor is computed using the RSM-based compact stress models. An individual TSV affecting this simulation point is visited and their stress contributions are added up. Once the stress computation at all points is finished, the von Mises stress and mobility maps for the full-chip are built.

The accuracy of linear superposition method compared with FEA simulations is already

---

**Algorithm 2:** RSM-Stress-Model based Full-Chip Stress and Reliability Analysis Flow.

---

```

input : TSV list  $T$ , material property distribution  $M$ , thermal map  $TM$ 
output: stress map, von Mises stress map, carrier mobility maps
1 for each TSV  $t$  in  $T$  do
2    $t.m \leftarrow \text{AssignMaterialProperty}(M, TM)$ 
3    $it \leftarrow \text{FindStressInfluenceZone}(t)$ 
4   for each point  $it'$  in  $it$  do
5      $it'.TSV \leftarrow t$ 
6   end
7 end
8 for each simulation point  $r$  do
9   if  $r.TSV \neq \emptyset$  then
10    for each  $t \in r.TSV$  do
11       $dt \leftarrow \text{distance}(t, r)$ 
12       $S_{cyl}(t) \leftarrow \text{ComputeStressRSM}(dt, t.m)$ 
13       $\theta(t) \leftarrow \text{GetAngle}(\text{line } tr, x\text{-axis})$ 
14       $Q(t) \leftarrow \text{SetConversionMatrix}(\theta_t, \theta_p, \theta_m)$ 
15       $S_{cart}(t) \leftarrow Q(t)S_{cyl}(t)Q(t)^T$ 
16       $r.S_{cart} \leftarrow r.S_{cart} + S_{cart}(t)$ 
17    end
18  end
19   $\text{vonMises}(r) \leftarrow \text{ComputeVonMises}(r.S_{cart})$ 
20   $\text{mobility}(r) \leftarrow \text{ComputeMobility}(r.S_{cart})$ 
21 end

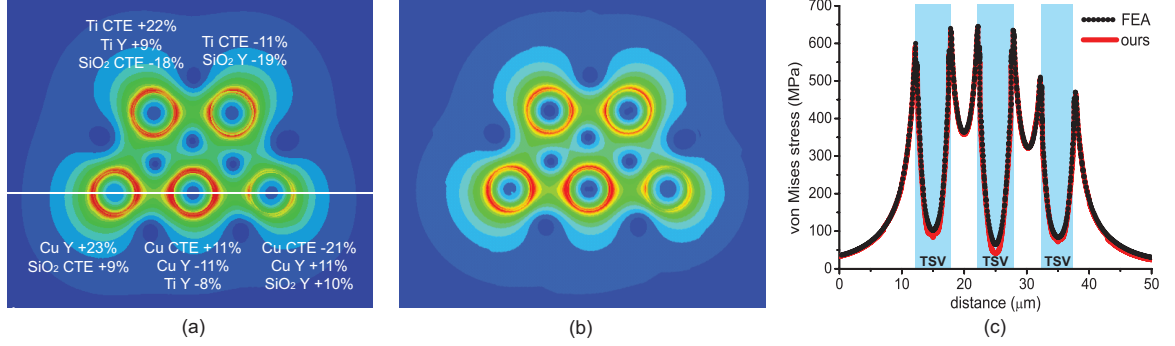
```

---

validated in [2]. The remaining question is whether this linear superposition method still holds under material property variations among TSVs. In order to answer this question, material properties are randomly assigned to each TSV and the analysis results are compared with FEA simulation results. Table 24 shows some of test cases. Results show that large errors occur inside TSV and TSV edge (= TSV/barrier interface). In general, the most critical region for mechanical reliability is at the interface between different materials. Thus, the TSV edge is important in this case. Although the maximum error at the TSV edge is as high as -18.6MPa, the %error is only -4.3%, which is acceptable for a fast full-chip stress analysis. The von Mises stress comparison of FEA and the proposed method is shown in Figure 33. This structure contains five TSVs with randomly assigned material properties. It clearly shows that the linear superposition method enhanced with RSM models produces highly accurate results within a fraction of runtime.

**Table 24:** Von Mises stress comparison between FEA and the proposed method. Error = ours - FEA.

# TSV	FEA		ours		max error (MPa)		
	# node	run time	# grid	run time	inside TSV	TSV edge	outside TSV
1	923K	1h34m	1M	2.5s	5.4	1.2	-1.7
2	1.1M	2h17m	1.2M	4.5s	-6.7	-11.4	-4.3
3	1.5M	2h53m	1.44M	6.2s	-13.2	-13.7	-3.2
5	2.2M	4h23m	2M	10s	-21.1	-18.6	-6.1



**Figure 33:** Von Mises stress comparison between FEA simulation and our method. TSV pitch is  $10\mu m$ . (a) FEA result. (b) Ours. (c) FEA vs. ours along the white line in (a).

## 4.6 Variation-Tolerant Full-Chip Design Methods

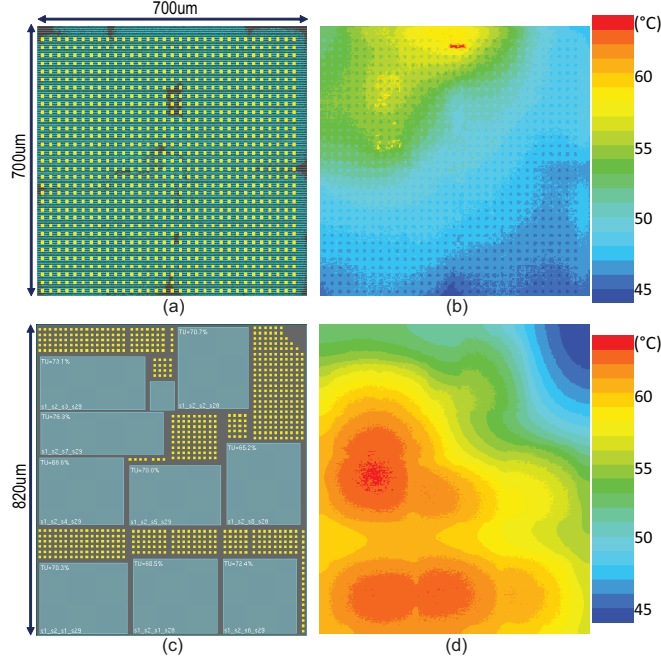
### 4.6.1 Simulation Settings

The RSM-model-based thermo-mechanical stress and reliability analysis flow is implemented in C++/STL. The impacts of material property variations on the full-chip thermo-mechanical stress and reliability as well as the 3D timing are examined in this section. In addition, design methods are explored to reduce mechanical reliability problems and performance degradations under material property variations.

For the simulations, two-die stack 3D IC designs are built with Nangate 45nm cell library. In addition, two different 3D IC design styles, i.e., gate-level and block-level designs, are employed as shown in Figure 34. Each 3D design contains 250K cells. The number of TSVs in gate-level and block-level designs are 1024 and 727, respectively. The baseline TSV material, pitch, and keep-out-zone (KOZ) size are copper,  $20\mu m$ , and  $1\mu m$ , respectively, unless otherwise specified.

1000 full-chip Monte Carlo simulations are performed for every test case, and both  $\sigma_{CTE}$

and  $\sigma_{Young's\ modulus}$  are set as 10% of the nominal value of each material, unless otherwise specified. It is also assumed that material properties follow the normal distribution. The maximum von Mises stress inside TSV material is monitored as a mechanical reliability metric for each simulation. In addition, the longest path delay (LPD) and total negative slack (TNS) are checked to examine the material property variation impact on the full-chip performance. The LPD and TNS are normalized to the *no variation* case ( $= [2]$ ).



**Figure 34:** Layouts and thermal maps of different 3D IC design styles. Yellow rectangles are TSVs. (a) Gate-level design with regular TSV placement (#TSV: 1024). (b) Thermal map of (a). (c) Block-level design (#TSV: 727). (d) Thermal map of (c).

#### 4.6.2 Robust Design with TSV Pitch Adjustment

In this section, the impact of TSV pitch on the full-chip mechanical reliability and performance is explored. Three different TSV pitches are used for both gate-level and block-level designs;  $15\mu m$ ,  $20\mu m$ , and  $25\mu m$ .

The mean of von Mises stress is the lowest with the smallest pitch as shown in Table 25. The von Mises stress is computed considering all stress tensor components. Inside the TSV array, there is destructive stress interference between tensile and compressive stress along the X- and Y-axis, respectively. This destructive stress interference becomes higher with a

smaller TSV pitch, and hence reduces von Mises stress level. However, the standard deviation of von Mises stress increases as the TSV pitch decreases. This is because the impact of stress variation in a TSV can easily propagate to nearby TSVs in smaller pitches. On the other hand, the mean and standard deviation of  $\sigma_{xx}$  magnitude increases monotonically with a smaller TSV pitch.

Nonetheless, the impact of TSV pitch on variation tolerance in mechanical reliability metrics is not significant. This is because mechanically instable spots in the TSV structure are TSV/barrier and barrier/liner interface in general, and in this region the TSV under consideration is the dominant stress contributor. Neighboring TSVs act as secondary stress contributors and their impact is limited.

As TSV pitch increases, variations in both LPD and TNS decrease significantly. The standard deviation of LPD and TNS decreases by 43.2% and 45.7%, respectively for gate-level designs, and by 42.9% and 64.6%, respectively for block-level designs as the TSV pitch increases from  $15\mu m$  to  $25\mu m$ . Therefore, this TSV pitch is a strong design knob to suppress performance variations under material property variations. However, increasing TSV pitch may lead to larger footprint area, hence a careful design decision needs to be made.

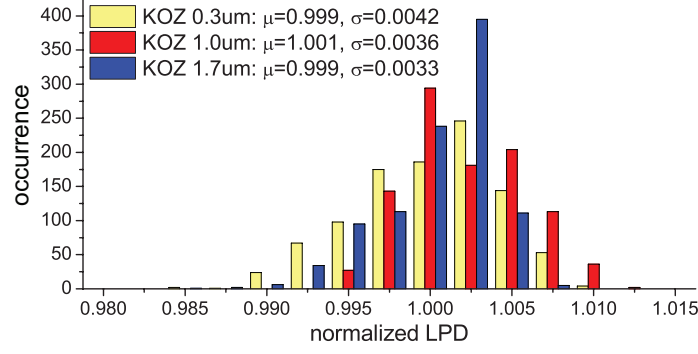
#### 4.6.3 Robust Design with TSV KOZ Size Adjustment

In this section, the impact of KOZ size on the full-chip performance under material property variations is examined. Three different KOZ sizes are employed;  $0.3\mu m$ ,  $1.0\mu m$ , and  $1.7\mu m$ . Because only the KOZ size is varied with fixed TSV locations, mechanical reliability metrics are not monitored. In addition, only gate-level designs with  $20\mu m$  TSV pitch are examined, since performance variations in block-level designs are much less than gate-level designs.

The standard deviation of LPD decreases monotonically with a larger TSV KOZ size as shown in Figure 35. Compared with the extreme  $0.3\mu m$  KOZ, the standard deviation of LPD decreases by 21.4% with  $1.7\mu m$  KOZ. Again, although increasing KOZ size could suppress the full-chip performance variation further, it might incur a larger footprint area and hence slower performance. Thus, the KOZ size needs to be determined carefully.

**Table 25:** Impact of TSV pitch on full-chip reliability and performance under material property variations. TSV material is Cu for all cases, and  $\sigma_{CTE}$  and  $\sigma_Y$  of each material are set as 10% of the nominal values.

TSV pitch ( $\mu m$ )	gate-level design						block-level design					
	von Mises (MPa)		$\sigma_{xx}$ (MPa)		normalized LPD		normalized TNS		von Mises (MPa)		$\sigma_{xx}$ (MPa)	
	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$	$\mu$	$\sigma$
15	585.3	24.8	234.3	8.6	1.000	0.0037	0.996	0.0302	576.5	18.8	230.6	8.7
20	600.0	22.3	230.7	8.3	1.001	0.0036	1.001	0.0296	593.1	18.4	224.6	8.4
25	597.9	22.8	228.3	8.2	1.000	0.0021	1.000	0.0164	588.3	17.9	214.3	8.3
											0.0004	0.0005
											0.999	0.0166
											0.993	0.0058



**Figure 35:** Impact of KOZ size on full-chip performance under variations.  $\sigma_{CTE}$  and  $\sigma_Y$  of each material are set as 10% of nominal values. Longest path delay is normalized to *no variation* case ( $= [2]$ ).

#### 4.6.4 Tungsten vs. Copper TSV on Full-Chip Robustness

The impact of TSV material on full-chip reliability and performance is studied in this section. The gate-level design with  $20\mu m$  TSV pitch is used. Since  $\sigma_{xx}$  of tungsten (W) TSV at the TSV and barrier interface is almost zero, only the maximum von Mises stress and LPD distributions are monitored.

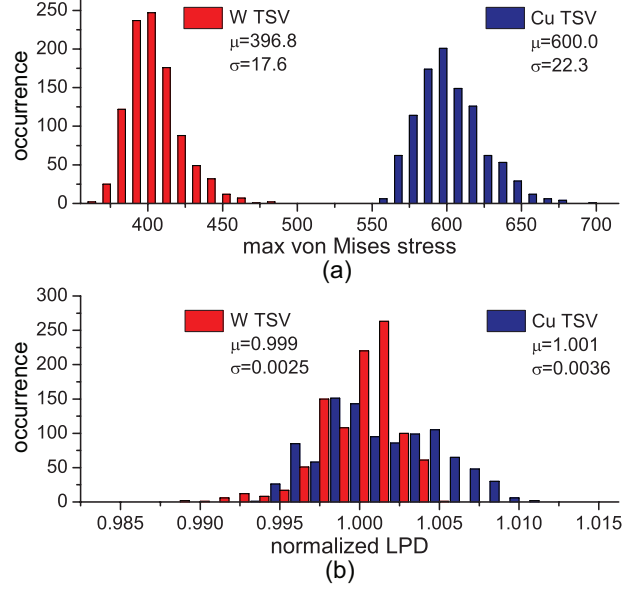
It is observed that the tungsten TSV shows narrower spread compared with copper TSV for both von Mises stress and LPD as shown in Figure 36. Moreover, the von Mises stress level is much lower than copper TSV. However, the yielding strength of tungsten is higher than copper. Thus, material yielding and interfacial delamination at the TSV/barrier interface for tungsten TSV are less of a concern than copper TSV.

However, the resistivity of tungsten ( $= 52.8n\Omega \cdot m$ ) is about 3.1X higher than copper ( $= 16.78n\Omega \cdot m$ ). This high resistive tungsten TSV may degrade both performance and 3D IR-drop. Therefore, the choice of TSV material needs to be determined carefully depending on applications.

#### 4.6.5 Design Style Choice: Block-level vs Gate-level 3D ICs

In this section, the impact of 3D IC design styles on the full-chip mechanical reliability and performance is investigated. The gate-level design has the potential of highest optimization because of the finest 3D partitioning granularity. In this work, TSVs are pre-placed uniformly across the die and then cells are placed [46]. On the other hand, in the block-level





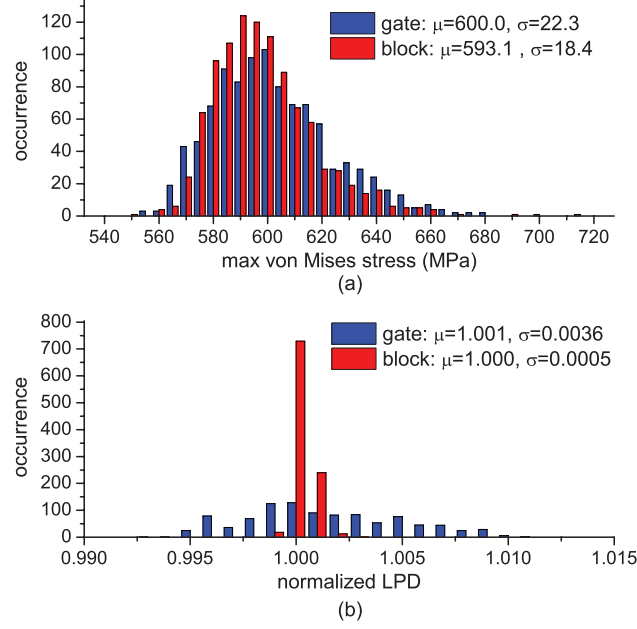
**Figure 36:** Impact of TSV material on full-chip reliability and performance under material property variations.  $\sigma_{CTE}$  and  $\sigma_Y$  of each material are set as 10% of the nominal value. (a) Maximum von Mises stress distribution. (b) Longest path delay (normalized to *no variation* case [2]) distribution.

design, 3D floorplanning is performed with existing highly-optimized 2D blocks, and TSVs are inserted into whitespace [52]. Thus, this block-level saves significant design cost compared with the gate-level design. Layouts and thermal maps (for temperature dependent CTE assignment) for both gate-level and block-level designs are shown in Figure 34.

Figure 37 shows distributions of maximum von Mises stress and LPD for both gate-level and block-level 3D designs. The wide spread of the maximum von Mises stress is observed in both design styles. Compared with the *no variation* case, the maximum von Mises stress increases by 75.3% and 72.6% for gate-level and block-level designs, respectively as shown in Table 26. This high stress can cause mechanical failures such as cracking.

It is also observed that the mean value of LPD does not deviate much from the *no variation* case. This is because the LPD is the path delay, not just a single cell delay, and hence some cells experience the mobility increase, while others undergo mobility degradation. However, in the gate-level design, we see that the LPD increases by up to 1.1% compared with the *no variation* case, which may not be ignorable in some applications. In the block-level design, since most cells inside blocks are far from TSV arrays, the full-chip performance

variation is almost negligible.



**Figure 37:** Impact of 3D IC design styles on full-chip reliability and performance under variations. (a) Max von Mises stress distribution. (b) Longest path delay (normalized to *no variation* case [2]) distribution.

**Table 26:** Maximum von Mises stress and longest path delay from 1000 Monte Carlo simulations. Numbers in parenthesis are % increase compared with the *no variation* ( $\sigma = 0$ ) case (= [2]).

design style	max von Mises stress (MPa)		max $\sigma_{xx}$ (MPa)	
	$\sigma = 0$	$\sigma = 10\%$	$\sigma = 0$	$\sigma = 10\%$
gate level	407.5	714.4 (+75.3%)	165.0	263.3 (+59.6%)
block level	399.8	690.1 (+72.6%)	168.1	258.6 (+53.8%)

#### 4.7 Summary

In this chapter, an accurate and fast full-chip stress and mechanical reliability analysis flow based on linear superposition principle of stress tensors is presented, which overcomes the limitation of FEA tools. This approach can be applicable to mechanical reliability aware placement optimization for 3D ICs. In addition, impacts of TSV surrounding structures such as a landing pad and a dielectric liner on stress fields and mechanical reliability in 3D ICs are examined. The TSV size and pitch, liner material and its thickness, and TSV placement are identified as key design parameters to reduce the mechanical reliability problems in TSV

based 3D ICs.

In addition, impacts of material property variations on the stress field, mechanical reliability, and performance in 3D ICs are studied. Stress models that consider CTE and Young's modulus of TSV, barrier, and liner materials are built using DOE and RSM. Based on these compact stress models, a full-chip thermo-mechanical stress and reliability analysis flow under material property variations is presented. The TSV pitch, KOZ size, and TSV material are key design knobs to reduce variations in mechanical reliability and performance under material property variations.

## CHAPTER V

### CHIP/PACKAGE MECHANICAL STRESS IMPACT ON 3D IC RELIABILITY AND MOBILITY VARIATIONS

Most previous works on the thermo-mechanical stress and reliability of TSV-based 3D ICs have been done separately in either chip or package domain. The impact of TSV-induced stress on device performance [53] and crack growth in TSV [54] were studied in the chip domain. As for the package domain, many works focused on the reliability of package bumps (= C4 bumps) [55]. Recently, authors in [15] showed the significant impact of package components on the chip domain stress. They proposed a stress exchange file to transfer the boundary conditions from package-level to silicon-level analysis.

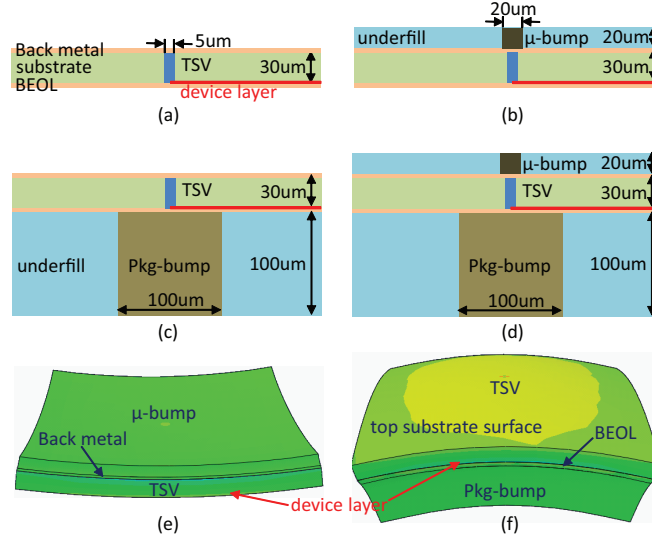
However, all of these approaches require FEA methods which are computationally expensive and hence infeasible for full-chip or -package analysis. To overcome the limitation of FEA method, linear superposition of stress tensors [5] and response surface method [48] were utilized. However, all of these were limited to the chip domain analysis.

The package bumps, underfill, and packaging substrate all add further mechanical stress to the 3D IC mounted above it in a non-trivial way. To accurately assess thermo-mechanical reliability problems and device performance variations in 3D IC/package systems, it is imperative to consider the interplay between the stress caused by the TSVs and the one by these packaging elements simultaneously. Moreover, to enable a chip/package co-design for better reliability and timing under the chip/package stress impact, a fast and accurate enough chip/package mechanical stress co-analysis tool is needed.

In this chapter, a full-chip/package-scale thermo-mechanical stress and reliability co-analysis flow as well as a design optimization methodology to reduce the mechanical reliability problems in TSV-based 3D ICs are presented. Additionally, the mobility and full-stack timing variations caused by the CTE mismatch among the materials are addressed in full-chip/package scale.

### 5.1 Motivation

It is first examined how various chip/package interconnect components interact and alter the thermo-mechanical stress distribution on the *device layer around TSV* caused by the CTE mismatch between TSV and substrate materials. First, TSV and substrate are only considered which most previous works studied. The same simulation structure used in [5] is employed as shown in Figure 38(a). Then, a  $\mu$ -bump and underfill layer *above* the substrate are added as shown in Figure 38(b). All structures undergo  $\Delta T = -250^\circ\text{C}$  of thermal load (annealing/reflow  $275^\circ\text{C} \rightarrow$  room temperature  $25^\circ\text{C}$ ). As Figure 39 shows, by adding the  $\mu$ -bump layer (= dotted red line), slightly more tensile (= positive) stress is observed than the TSV-only case (= solid black line). This is because  $\Delta\text{CTE}$  of  $\mu$ -bump and underfill is 24ppm/K, while that of TSV and substrate is 14.7 ppm/K, hence the deformation of the entire structure is largely determined by the  $\mu$ -bump and underfill layer. Since the top side of  $\mu$ -bump layer is free surface, the entire structure easily bends upward as all the elements shrink from the negative thermal load as shown in Figure 38(e). Thus, the materials on device layer stretch outward, which results in more tensile stress.

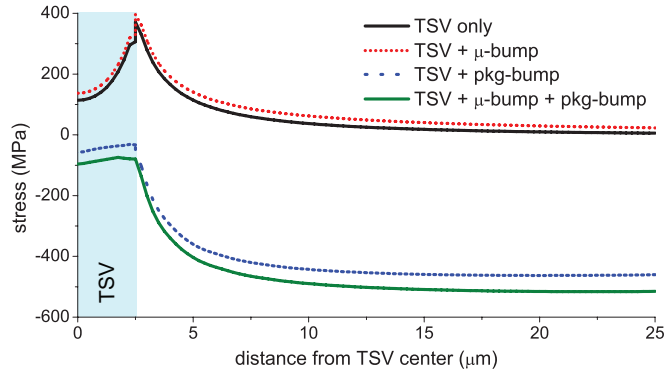


**Figure 38:** Impact of bumps and underfill on the stress of device layer (= red line). (a) TSV only [5] (b) TSV +  $\mu$ -bump (c) TSV + package-bump (d) TSV +  $\mu$ -bump + package-bump. (e) Deformed structure of (b). (f) Deformed structure of (c). Both (e) and (f) are drawn with 10X deformation scale factor.

On the other hand, if a package-bump (= C4 bump) layer *below* the substrate is added

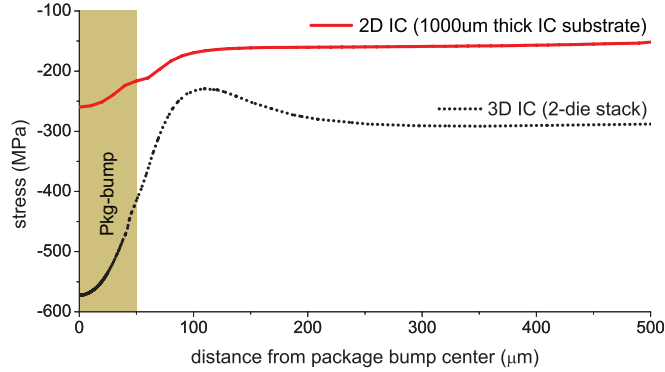
as shown in Figure 38(c), now the entire structure bends downward as shown in Figure 38(f) because package elements are shrinking more than chip elements. The  $\Delta\text{CTE}$  of package bump and underfill is 22ppm/K. This generates highly compressive (= negative) stress on the device layer. Comparing Figure 38(b) and Figure 38(c), it is observed that the bending direction depends on which layer shrinks more: in both cases, the bump layers shrink more than the silicon substrate.

Lastly, both bump layers are included as shown in Figure 38(d). In this case, the  $\Delta\text{CTE}$  is almost the same (24ppm/K on the top, 22ppm/K on the bottom). However, the overall structure bends down in a similar fashion as shown in Figure 38(f) because of the sheer volume of package bump layer (= shrinking more than the  $\mu$ -bump layer). This in turn causes compressive stress in the device layer. However, the magnitude is slightly *more* (= solid green line in Figure 39) than the package-bump layer only case (= dotted blue line). One might expect the overall compressive stress would be less because the  $\mu$ -bump layer tries to bend upward while the package-bump layer tries to bend downward (= canceling effect). However, this additive effect is because the  $\mu$ -bump layer eventually bends down and adds more compressive stress to the device layer. Note that the bending direction of the  $\mu$ -bump layer is affected by adjacent layers. Since now the deformation of the entire structure is dominated by the package-bump layer, the flexible underfill material in the  $\mu$ -bump layer easily bends downward. These basic simulations clearly show the importance of considering package element impact on the chip-domain stress distribution.



**Figure 39:** Impact of package components on the stress ( $\sigma_{rr}$ ) around TSV on device layer (FEA results).

Figure 40 shows the stress contributions of package bump and underfill layer to the chips (2D vs. 3D) mounted on it. For the 3D IC/package structure, a two-die stack chip/package structure is built similar to Figure 41(a) excluding TSV and  $\mu$ -bump. This is to examine the impact of package-bump solely. The bottom die (= die0) is thinned, and the device layer of this thin die is examined. One 2D IC/package structure is also created, where a single un-thinned die of 1000  $\mu m$  thickness is employed. The device layer of this un-thinned die is monitored. The same thermal load ( $\Delta T = -250^\circ C$ ) is applied for both cases. As shown in Figure 40, the 3D IC experiences more severe compressive stress than the 2D IC case. The main reason is the thickness and the flexibility of the die that is monitored. Even though the thickness of the entire structure is thicker in 3D IC, the thin die (30  $\mu m$  thick) and the underfill material above the thin die is much more flexible than the un-thinned substrate in 2D IC. Thus, this thin die is highly affected by the package-bump underneath it. This indicates that the impact of package-bump is more significant in 3D IC.



**Figure 40:** Comparison of impact of package-bump on the device layer stress ( $\sigma_{rr}$ ) between 2D IC and 3D IC (2-die stack) (FEA results).

## 5.2 3D IC/Package Stress Modeling

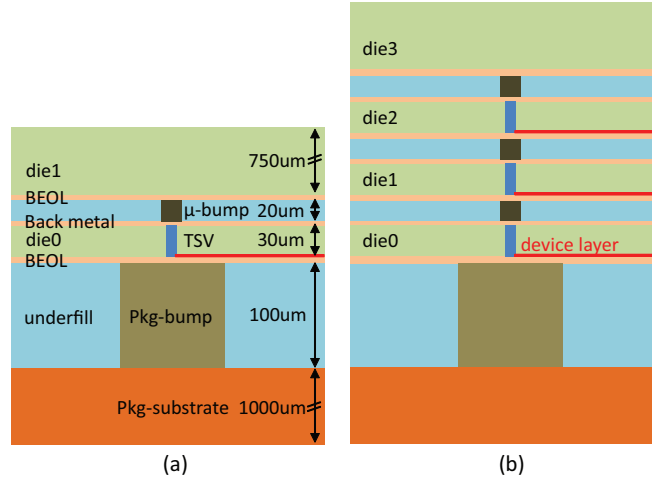
The von Mises yield criterion [43] is used as a mechanical reliability metric for TSVs. However, no specific threshold value for the von Mises criterion is used, since it is greatly affected by fabrication process. Von Mises stress is computed using Equation (12).

$$\sigma_v = \sqrt{\frac{(\sigma_{xx} - \sigma_{yy})^2 + (\sigma_{yy} - \sigma_{zz})^2 + (\sigma_{zz} - \sigma_{xx})^2 + 6(\sigma_{xy}^2 + \sigma_{yz}^2 + \sigma_{zx}^2)}{2}} \quad (12)$$

### 5.2.1 Chip/Package Co-Simulation Structure

The simulation structure is shown in Figure 41, where the dimensions of baseline simulation structures are based on the fabricated and/or published data [15, 24]. In this work, the stress distribution on device layer for each die is examined shown in red lines in Figure 41. The baseline TSV diameter, height, landing pad size, Cu diffusion barrier thickness, and dielectric liner thickness are  $5\mu m$ ,  $30\mu m$ ,  $6\mu m$ ,  $50nm$ , and  $125nm$ , respectively. Ti and  $SiO_2$  are used as Cu diffusion barrier and liner materials. Also, the diameter/height of  $\mu$ -bump and package-bump are  $20\mu m$  and  $100\mu m$ , respectively, unless otherwise specified. Material properties used for the experiments are as follows: CTE (ppm/K) / Young's modulus (GPa) for Cu = (17/110), Si = (2.3/188),  $SiO_2$  = (0.5/71), Ti = (8.6/116), package-bump (SnCu) = (22/44.4),  $\mu$ -bump ( $Sn_{97}Ag_3$ ) = (20/26.2), underfill = (44/5.6), package substrate (FR-4) = (17.6/19.7).

A FEA simulation tool ABAQUS [37] is employed to perform experiments, and all materials are assumed to be linear elastic and isotropic. The entire structure undergoes  $\Delta T = -250^\circ C$  of thermal load (annealing/reflow  $275^\circ C \rightarrow$  room temperature  $25^\circ C$ ) to represent a fabrication process. In addition, all materials are assumed to be stress free at the annealing/reflow temperature.



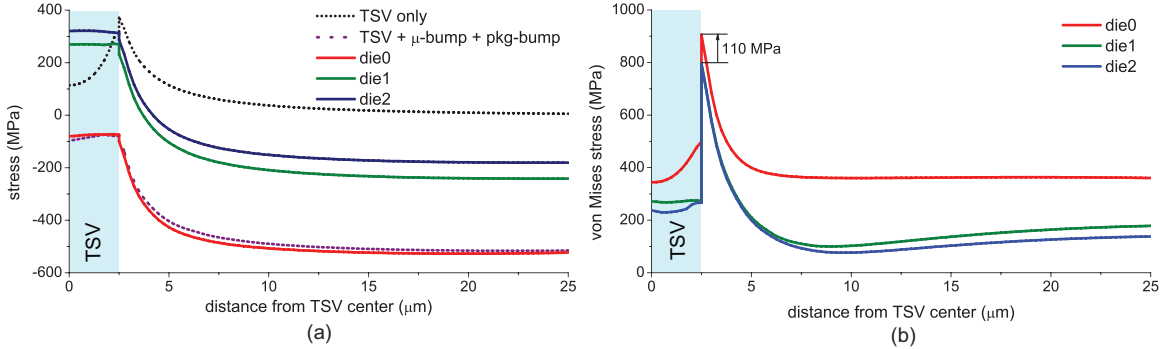
**Figure 41:** Side view of baseline chip/package simulation structures. (a) 2-die stack (b) 4-die stack.



### 5.2.2 Impact of Die Stacking

Previous works on the full-chip thermo-mechanical stress analysis used the same stress pattern for different dies in a multiple-die stack [5, 53]. In this section, the impact of die stacking on the thermo-mechanical stress distribution on the device layer across strata is examined. A four-die stack structure is employed for this purpose. In addition, only one TSV,  $\mu$ -bump, and package-bump are used for each die or layer, respectively, and their center locations are aligned as shown in Figure 41.

First of all, the stress level, the extent of compression or tension, differs significantly across dies as shown in Figure 42(a). The overall stress trend remains similar: the stress is highest at TSV edge and decays then saturates as distance increases from the TSV center. However, the bottom-most die (= die0, solid red line), which is closest to the package-bump layer, shows most compressive stress among three dies containing TSV. This is because the impact of package-bump is most significant in die0 due to their proximity.



**Figure 42:** Impact of die stacking on device layer stress (FEA results). (a)  $\sigma_{rr}$  stress on device layer in each die in 4-die stack. (b) Von Mises stress in each die in a 4-die stack.

In addition, as we go to the upper dies, the stress level becomes closer to the case considering only TSV and substrate. It is also observed that the stress curve of die0 is very close to the case of TSV +  $\mu$ -bump + package-bump (= dotted purple line), which does not contain the package substrate and un-thinned top die shown in Figure 38(d). This also indicates that the stress level in die0 is mostly determined by package-bump. The stress distribution in die3 (un-thinned top die without TSVs) is almost flat ( $-110 \pm 5$  MPa). Since die3 does not contain any TSVs, there is no local von Mises stress peak (= dangerous

region) caused by TSVs. Thus, only the dies containing TSVs are considered in this work.

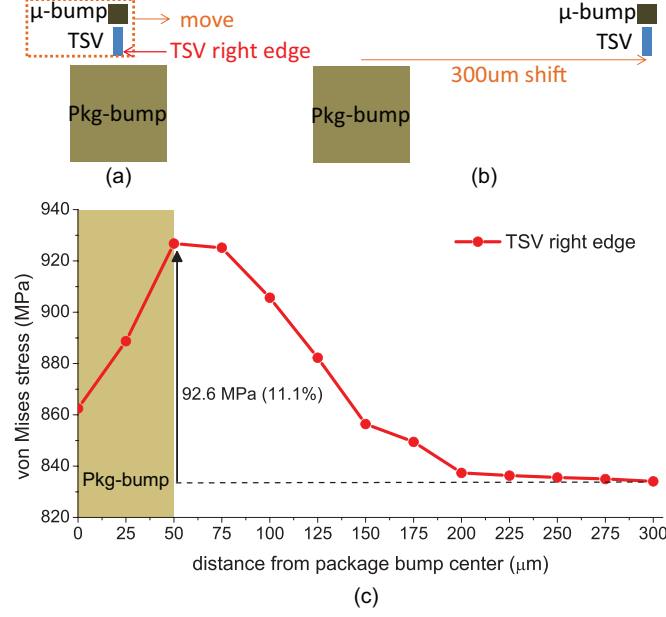
Moreover, the mechanical reliability problem is most severe in die0 as shown in Figure 42(b). The maximum von Mises stress at TSV edge in die0 is about 110MPa higher than the upper two dies. This is again mostly due to the package-bump that induces large deformation at the nearest die.

### 5.2.3 Impact of TSV and Bump Alignment

Now the impact of alignment between TSV,  $\mu$ -bump, and package-bump on the mechanical reliability of TSVs is studied. First, the impact of relative position between TSV/ $\mu$ -bump and package-bump is examined. A two-die stack structure in which center locations of TSV,  $\mu$ -bump, and package-bump are aligned is used as shown in Figure 43(a). Then both TSV and  $\mu$ -bump are shifted together from the package-bump center with a  $25\mu m$  step, and the von Mises stress at the right edge of TSV is monitored.

The von Mises stress is maximum around the package-bump edge region and then decreases and saturates as the distance increases as shown in Figure 43(c). The difference between minimum and maximum is as high as 11.1%. The highest stress gradient occurs around the package-bump edge which results in the highest deformation of the structure near this region. Hence, this higher deformation causes more severe mechanical reliability problem in TSV. Interestingly, the von Mises stress decreases near the package-bump center. This is because the material around this area is the same (= package-bump material), hence its deformation is relatively smaller than the edge which is the interface between two different materials.

In addition, it is examined whether the relative position between  $\mu$ -bump and TSV/package-bump affects the mechanical reliability of TSV. The location of TSV and package-bump are fixed and their centers are aligned, then only  $\mu$ -bump is moved with a  $5\mu m$  step up to  $30\mu m$ . The von Mises stress at the TSV edges is monitored. The similar trend is observed as before. However, the difference between minimum and maximum is only 6.5MPa (0.8%), which is negligible. Thus, the relative position between TSV and package-bump is identified as a critical factor that affects the mechanical reliability of TSV.



**Figure 43:** Impact of relative position between TSV/ $\mu$ -bump and package-bump on von Mises stress. (a) Initial position. (b) Final position where TSV/ $\mu$ -bump are shifted by  $300\mu m$  from package bump center. (c) Von Mises stress at TSV edge along the distance between TSV/ $\mu$ -bump and package-bump (FEA results).

### 5.3 Mobility Variation Modeling

#### 5.3.1 Need for True 3D Chip/Package Stress Model

The analytical 2D radial stress model, known as *Lamé* stress solution, was employed to address the TSV thermo-mechanical stress. This 2D plane solution assumes an infinitely long TSV embedded in an infinite silicon substrate and provides stress distribution in silicon substrate region, which can be expressed as follows [33]:

$$\begin{aligned}\sigma_{rr}^{Si} &= -\sigma_{\theta\theta}^{Si} = -\frac{E\Delta\alpha\Delta T}{2} \left( \frac{D_{TSV}}{2r} \right)^2 \\ \sigma_{zz}^{Si} &= \sigma_{rz}^{Si} = \sigma_{\theta z}^{Si} = \sigma_{r\theta}^{Si} = 0\end{aligned}\tag{13}$$

where  $\sigma^{Si}$  is stress in silicon substrate,  $E$  is Young's modulus,  $\Delta\alpha$  is mismatch in CTE,  $\Delta T$  is differential thermal load,  $r$  is the distance from TSV center, and  $D_{TSV}$  is TSV diameter.

Authors in work [11] used this 2D analytical solution to assess the impact of TSV-induced stress on the mobility variation and full-chip timing. However, in [11] only  $\sigma_{rr}$  stress term was considered while all other eight stress tensor elements were set to zero. When only one normal stress component is considered, we call this uniaxial stress. However, stress is

biaxial in nature in an elastic object as Equation (13) indicates: there exist two non-zero normal stress components, i.e.,  $\sigma_{rr}$  and  $\sigma_{\theta\theta}$ . Since the mobility variation depends on the piezoresistive effect due to stress, the mobility variation pattern may change depending on the choice of stress mode.

Although this closed-form formula is easy to handle, this 2D solution is only applicable to the structure with TSV and substrate only, hence it is inappropriate for the realistic TSV structure with a Cu diffusion barrier and a dielectric liner. In addition, a huge stress magnitude discrepancy was observed around TSV edge on the device layer between the 2D stress model and the 3D FEA simulations [5]. This is simply because a 3D TSV structure cannot be correctly modeled by the 2D plane solution due to the change in boundary conditions, especially near the top and bottom of the structure. Moreover, packaging elements and die-stacking affect stress distribution on each device layer differently. Therefore, if the 3D stress tensors, i.e., non-zero nine stress components, as well as packaging elements are considered, the mobility variation pattern can be significantly different from 2D stress cases.

### 5.3.2 Piezoresistivity

In semiconductors, changes in inter-atomic spacing resulting from strain affect the bandgaps, making it easier or harder for electrons—depending on the material and strain—to be raised into the conduction band. This results in a change in resistivity of the semiconductor, which also can be translated to a change in mobility as follows [1]:

$$\begin{aligned} \frac{\Delta R}{R} = -\frac{\Delta\mu}{\mu} = & [\pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy}] \cos^2\phi \\ & + [\pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy}] \sin^2\phi \\ & + \pi_{12}\sigma_{zz} + \pi'_{44}\sigma_{xy}\sin 2\phi \end{aligned} \quad (14)$$

where  $\sigma_{ij}$  is the stress in the silicon substrate in Cartesian coordinate system, and  $\phi$  is an angle between the wafer orientation and the transistor channel.

In this work, the (100) Si wafer with reference axes of  $[110]$ ,  $[\bar{1}10]$ , and  $[001]$  is assumed. It is also assumed that the transistor channel direction and the x-axis ( $[110]$ ) are identical. In this setup,  $\pi'_{ij}$  is the piezoresistivity coefficient defined along the reference axes of (100)

Si wafer listed in Table 27:

$$\begin{aligned}\pi'_{11} &= \frac{\pi_{11} + \pi_{12} + \pi_{44}}{2} \\ \pi'_{12} &= \frac{\pi_{11} + \pi_{12} - \pi_{44}}{2} \\ \pi'_{44} &= \pi_{11} - \pi_{12}\end{aligned}$$

Note that the piezoresistivity coefficients in Table 27 were obtained under 1.5GPa biaxial strain [3]. Thus, the mobility analysis results can provide an accurate assessment of stress impact on device performance and full-chip timing variations in deep sub-micron technologies. Many previous works [11, 53] used piezoresistivity coefficients for lightly doped n- and p-type silicon without any strain. From the mobility simulations, the case with piezoresistivity coefficients without strain shows up to 46% more mobility variations than the case with strain. In the latter case, the silicon is already highly stress engineered, hence the impact of TSV stress on the mobility variation reduces.

**Table 27:** Piezoresistive coefficient ( $TPa^{-1}$ ) in (100) Si wafer [3].

type	$\pi_{11}$	$\pi_{12}$	$\pi_{44}$	$\pi'_{11}$	$\pi'_{12}$	$\pi'_{44}$
N-type Si	-650	330	-120	-220	-100	-980
P-type Si	-40	30	970	480	-490	-70

### 5.3.3 Mobility Variation: 2D vs. 3D Stress

In this section, the impact of different stress cases on the mobility variation around a single TSV is examined. To utilize Equation (14), stress tensors need to be converted from cylindrical coordinate system ( $S_{r\theta z}$ ) to Cartesian coordinate system ( $S_{xyz}$ ).

$$S_{xyz} = \begin{bmatrix} \sigma_{xx} & \sigma_{xy} & \sigma_{xz} \\ \sigma_{yx} & \sigma_{yy} & \sigma_{yz} \\ \sigma_{zx} & \sigma_{zy} & \sigma_{zz} \end{bmatrix} S_{r\theta z} = \begin{bmatrix} \sigma_{rr} & \sigma_{r\theta} & \sigma_{rz} \\ \sigma_{\theta r} & \sigma_{\theta\theta} & \sigma_{\theta z} \\ \sigma_{zr} & \sigma_{z\theta} & \sigma_{zz} \end{bmatrix}$$

The transform matrix Q is the form:

$$Q = \begin{bmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

where  $\theta$  is the angle between the x-axis and a line from the origin to the center of a transistor channel. A stress tensor in a cylindrical coordinate system can be converted to a Cartesian coordinate system using conversion matrices:  $S_{xyz} = QS_{r\theta z}Q^T$ .

Now the impact of different stress cases on the mobility variation pattern is examined. First, stress tensor components in Equation (14) are shown. Then, the mobility variation formula is derived for each case. It is assumed that the x-axis and the transistor channel direction are identical ( $\phi = 0$ ).

- **2D uniaxial stress:**  $\sigma_{rr} \neq 0$ , all other stress terms = 0

$$\begin{aligned}\sigma_{xx} &= \sigma_{rr}\cos^2\theta, \sigma_{yy} = \sigma_{rr}\sin^2\theta, \sigma_{zz} = 0 \\ -\Delta\mu/\mu &= \pi'_{11}\sigma_{rr}\cos^2\theta + \pi'_{12}\sigma_{rr}\sin^2\theta\end{aligned}\tag{15}$$

- **2D biaxial stress:**  $\sigma_{rr} = -\sigma_{\theta\theta} \neq 0$ , all other stress terms = 0

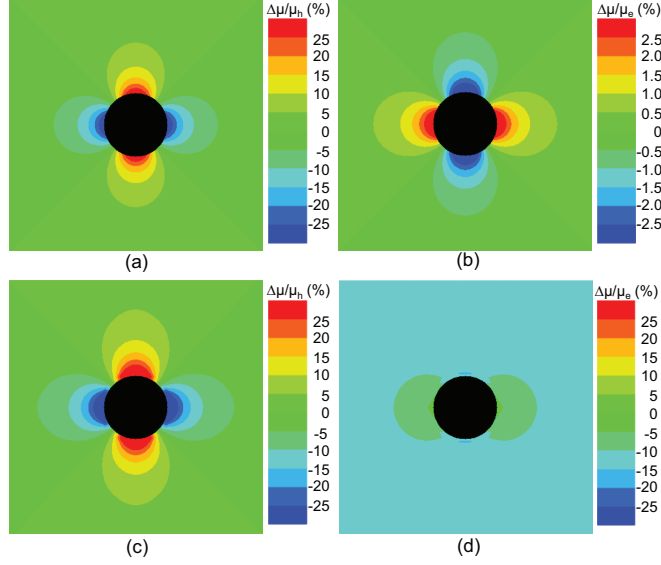
$$\begin{aligned}\sigma_{xx} &= -\sigma_{yy} = \sigma_{rr}\cos 2\theta, \sigma_{zz} = 0 \\ -\Delta\mu/\mu &= \pi'_{11}\sigma_{rr}\cos 2\theta - \pi'_{12}\sigma_{rr}\cos 2\theta = \pi_{44}\sigma_{rr}\cos 2\theta\end{aligned}\tag{16}$$

- **3D stress:** all stress tensor components  $\neq 0$

$$\begin{aligned}\sigma_{xx} &= \sigma_{rr}\cos^2\theta + \sigma_{\theta\theta}\sin^2\theta - \sigma_{r\theta}\sin 2\theta \\ \sigma_{yy} &= \sigma_{rr}\sin^2\theta + \sigma_{\theta\theta}\cos^2\theta + \sigma_{r\theta}\sin 2\theta \\ \sigma_{zz} &\neq 0 \\ -\Delta\mu/\mu &= \pi'_{11}\sigma_{xx} + \pi'_{12}\sigma_{yy} + \pi_{12}\sigma_{zz}\end{aligned}\tag{17}$$

It is clear from above expressions that the trend of mobility variation is different between these stress cases. Mobility variation maps around a single TSV for the 2D biaxial stress (*2D biaxial*) and the 3D stress with package components (*3D wPkg*) are shown in Figure 44. A significant difference in the electron mobility variation maps is observed, which is discussed in detail in Section 5.5.

Comparing both 2D stress cases, it is observed that the electron mobility in the 2D uniaxial stress (*2D uniaxial*) improves regardless of angle  $\theta$ , since both  $\pi'_{11}$  and  $\pi'_{12}$  are

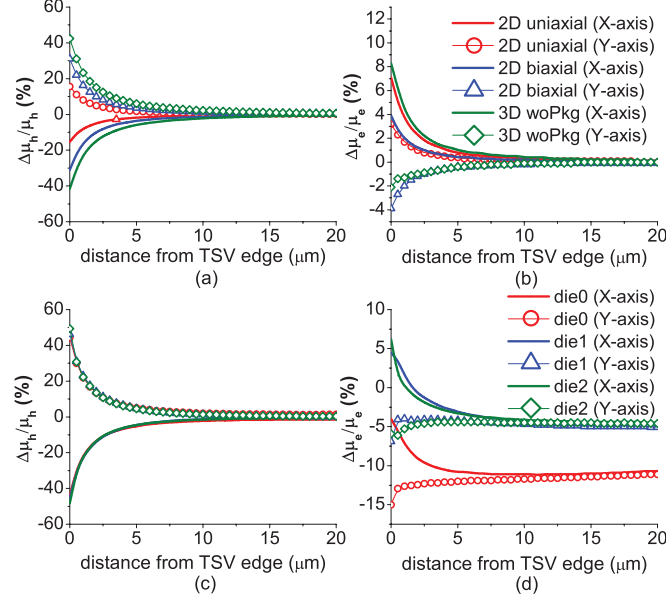


**Figure 44:** Mobility variation map around a single TSV. (a) Hole mobility (2D biaxial stress). (b) Electron mobility (2D biaxial stress). (c) Hole mobility in die0 in 4-die stack (3D stress with package components). (d) Electron mobility in die0 in 4-die stack (3D stress with package components). For both (c) and (d) TSV,  $\mu$ -bump, and package-bump are vertically aligned.

negative for N-type silicon and  $\sigma_{rr}\cos^2\theta$  and  $\sigma_{\theta\theta}\sin^2\theta$  terms are non-negative. On the other hand, the sign of electron mobility variation in the *2D biaxial* case depends on  $\theta$ , which is shown in Figure 45(b). We also observe that the *2D uniaxial* case underestimates the hole mobility variation range compared with the *2D biaxial* case. Thus, using *2D uniaxial* model in [11] may result in erroneous results.

As for the 3D stress without package components case (*3D woPkg*) shown in Figure 45(a) and 45(b), the hole mobility variation range is larger than the *2D biaxial* case. In addition, the electron mobility variation is not symmetric along the x-axis and the y-axis unlike the *2D biaxial* case. This is largely due to the non-zero  $\sigma_{zz}$  term. Note that in cases of *2D uniaxial*, *2D biaxial*, and *3D woPkg*, stress tensors are assumed to be identical across tiers, hence there is no difference in mobility variations in different dies in the 3D stack.

As package components are included, the electron mobility variation differs across the stack as shown in Figure 45(d). This is mainly due to the large compressive stress generated by the package-bump. This effect is most significant in die0, which is closest to the package-bump layer shown in Figure 41. More details are discussed in Section 5.5.



**Figure 45:** Mobility variation range of a single TSV with different stress cases. Mobility variation numbers are collected along the x-axis and the y-axis from a TSV center on device layers. (a) Hole mobility under 2D and 3D stress without package components. (b) Electron mobility under 2D and 3D stress without package components. (c) Hole mobility under 3D stress with package components in 4-die stack. (d) Electron mobility under 3D stress with package components in 4-die stack.

#### 5.4 Handling Full-Chip/Package: The LVLS Method

FEA simulation for multiple TSVs,  $\mu$ -bumps, and package-bumps require huge computing resources and time, thus it is not feasible for a full-system-scale analysis. In this section, a chip/package thermo-mechanical stress co-analysis flow in full-chip/package scale is presented. The principle of lateral and vertical linear superposition of stress tensors from individual TSVs,  $\mu$ -bumps, and package-bumps is utilized to enable a full-system-level analysis. The proposed approach is validated against FEA simulation results. Based on the linear superposition method, full-chip stress maps are built. Then compute the von Mises yield metric to assess the mechanical reliability problems in TSV-based 3D ICs.

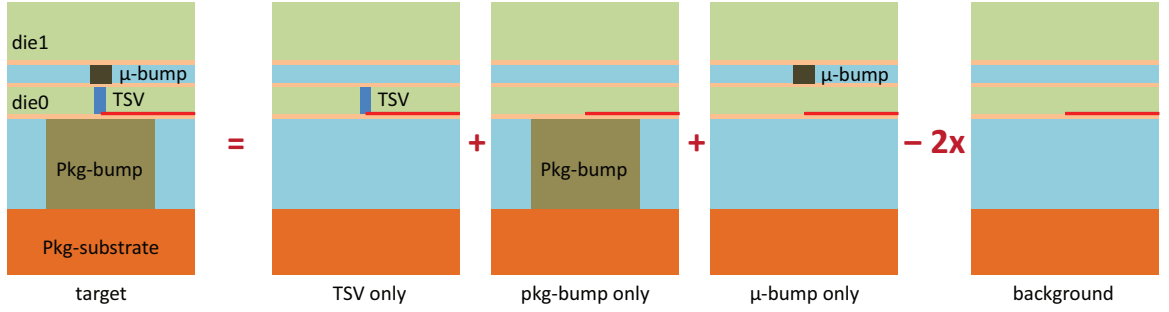
##### 5.4.1 Lateral and Vertical Linear Superposition

In [5], authors used the principle of linear superposition of stress tensors to perform a full-chip stress and reliability analysis considering many TSVs. In that case, all stress contributors (= TSVs) are on the same layer, hence we call this lateral linear superposition.



However, as the impact of  $\mu$ -bump and package-bump is considered, which are not in the same layer where TSVs are located, this lateral linear superposition cannot be used alone. Fortunately, the principle of linear superposition is not limited to 2D plane, but applicable to any linearly elastic structures including 3D structures.

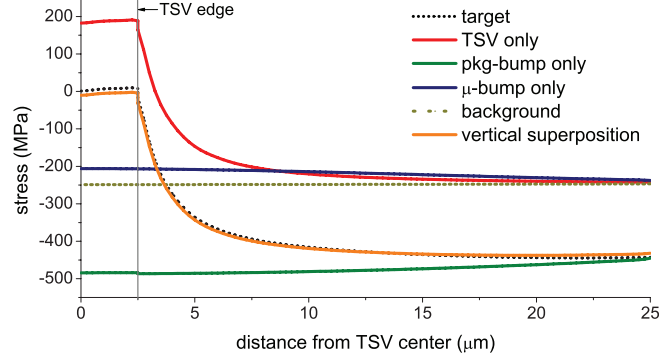
Figure 46 illustrates the vertical linear superposition method, which enables us to consider the stress induced by elements which are not in the same layer. First the target structure is decomposed into four separate structures: TSV only, package-bump only,  $\mu$ -bump only, and background which does not contain TSV and bumps. Next, stress tensors affected by each interconnect element are obtained along the red line on device layer from FEA simulations. Then, the stress tensors from TSV only, package-bump only, and  $\mu$ -bump only structures are added up, and twice the magnitude of the background stress tensors is subtracted since this background stress is already included in previous three structures. If the point under consideration is affected by  $n$  components, then it is required to subtract  $n-1$  times the background stress.



**Figure 46:** Illustration of vertical linear superposition with a 2-die stack structure. Stress is extracted along the red line on device layer from each structure using FEA tool.

Figure 47 shows the stress distributions from each structure as well as the stress obtained by the vertical linear superposition. It is observed that  $\mu$ -bump induces more tensile stress than background and package-bump generates much more compressive stress than background, which is discussed in Section 5.1. In addition, even without interconnect elements (= background), device layer is in compression due to the shrinking of the underfill material which has the highest CTE (= 44ppm/K) among all materials in the simulation structure. Most importantly, the vertical linear superposition method matches well with the

target stress distribution. Although the maximum error (11MPa) occurs inside TSV, this is inevitable since the direct vertical interaction between TSV,  $\mu$ -bump, and package-bump is ignored by decomposing the structure. Nonetheless, this error is acceptable for a fast full-system-scale analysis.



**Figure 47:** Vertical linear superposition of  $\sigma_{rr}$  stress in a 2-die stack shown in Figure 46

To obtain the stress tensor at a point affected by multiple TSVs,  $\mu$ -bumps, and package-bumps, both lateral and vertical linear superposition (LVLS) are applied as follows:

$$S = \sum_{i=1}^{n_{TSV}} S_{TSV_i} + \sum_{j=1}^{n_{\mu B}} S_{\mu B_j} + \sum_{k=1}^{n_{pkgB}} S_{pkgB_k} - (n_{TSV} + n_{\mu B} + n_{pkgB} - 1) \times S_{bg} \quad (18)$$

where,  $S$  is the total stress at the point under consideration and  $S_{TSV_i}$ ,  $S_{\mu B_j}$ , and  $S_{pkgB_k}$  are individual stress tensor at this point due to  $i^{th}$  TSV,  $j^{th}$   $\mu$ -bump, and  $k^{th}$  package-bump, respectively.  $S_{bg}$  indicates the background stress at that point.

#### 5.4.2 Full-Chip/Package Stress Analysis Flow

In this section, a full-chip/package stress analysis flow based on the LVLS method shown in Algorithm 3 is briefly explained. First, a stress library is built from FEA simulations. This library contains stress tensors along an arbitrary radial line on the device layer induced by each interconnect element, i.e., TSV,  $\mu$ -bump, and package-bump, separately. Given locations of TSVs,  $\mu$ -bumps, and package-bumps, a stress influence zone is found for each element. Beyond this stress influence zone of each interconnect element, the stress induced by the element under consideration is negligible [5]. In this work, five times the diameter

---

**Algorithm 3:** Full-Chip/Package Stress and Reliability Analysis Flow (LVLS)

---

```
input : TSV list  $T$ , pkg-bump list  $P$ ,  $\mu$ -bump list  $M$ , stress library
output: stress map, von Mises stress map, carrier mobility map
1 for each TSV  $t$ , pkg-bump  $p$ , and  $\mu$ -bump  $m$  in  $T$ ,  $P$ , and  $M$  do
2    $(it, ip, im) \leftarrow \text{FindStressInfluenceZone}(t, p, m)$ 
3   for each point  $it'$ ,  $ip'$ , and  $im'$  in  $it$ ,  $ip$ , and  $im$  do
4      $it'.TSV \leftarrow it$ 
5      $ip'.pkg\text{-}bump \leftarrow ip$ 
6      $im'.\mu\text{-}bump \leftarrow im$ 
7   end
8 end
9 for each simulation point  $r$  do
10  if  $r.TSV \neq \emptyset$  or  $r.pkg\text{-}bump \neq \emptyset$  or  $r.\mu\text{-}bump \neq \emptyset$  then
11    for each  $(t, p, m) \in (r.TSV, r.pkg\text{-}bump, r.\mu\text{-}bump)$  do
12       $(dt, dp, dm) \leftarrow \text{distance}(t, p, m, r)$ 
13       $S_{cyl}(t, p, m) \leftarrow \text{GetStressTensor}(dt, dp, dm)$ 
14       $S_{cyl}(t, p, m) \leftarrow S_{cyl}(t, p, m) - BGstress$ 
15       $\theta(t, p, m) \leftarrow \text{GetAngle}(\text{line } tr, pr, mr, x\text{-axis})$ 
16       $Q(t, p, m) \leftarrow \text{SetConversionMatrix}(\theta_t, \theta_p, \theta_m)$ 
17       $S_{Cart}(t, p, m) \leftarrow Q(t, p, m)S_{cyl}(t, p, m)Q(t, p, m)^T$ 
18       $r.S_{Cart} \leftarrow r.S_{Cart} + S_{Cart}(t, p, m)$ 
19    end
20  end
21   $r.S_{Cart} \leftarrow r.S_{Cart} + BGstress$ 
22   $vonMises(r) \leftarrow \text{ComputeVonMises}(r.S_{Cart})$ 
23   $mobility(r) \leftarrow \text{ComputeMobility}(r.S_{Cart})$ 
24 end
```

---

of each component is used as a stress influence zone, which is determined by FEA simulations. Then, each grid point is associated with all the interconnect elements whose stress influence zone overlaps with the point. Next, the LVLS method is applied at the point under consideration to obtain the stress tensor induced by every component found in the association step. In this step, the coordinate conversion matrices are used to obtain stress tensors in the Cartesian coordinate system. Finally, the von Mises stress value is computed using Equation (12) to assess the mechanical reliability problem in TSVs and mobility maps using Equation (14).

#### 5.4.3 Validation of LVLS

In this section, the LVLS method is validated against FEA simulations by varying the number of TSVs,  $\mu$ -bumps, and package-bumps as well as their arrangement. The minimum pitch of TSV,  $\mu$ -bump, and package-bump are set as  $10\mu m$ ,  $20\mu m$ , and  $200\mu m$  for all test

cases. Stress tensors along the radial line on device layer induced by each interconnect element (stress tensor library) are obtained through FEA simulations with  $0.25\mu m$  interval. In the LVLS method, the simulation area is divided into uniform array style grid with  $0.1\mu m$  pitch. If the stress tensor at the grid point under consideration is not obtainable directly from the stress library, the stress tensor is computed using linear interpolation with adjacent stress tensors in the library.

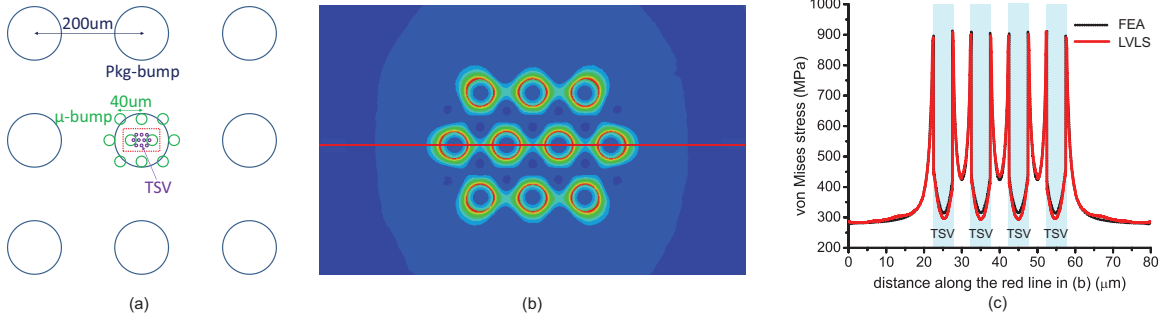
Some of our comparisons in die0 in a four-die stack are listed in Table 28. The die0 shows the largest errors among three dies containing TSVs due to its proximity to package-bumps. In addition, only the cases with the minimum pitches for each component are listed, which again shows maximum errors. First, a huge run time reduction is observed in the LVLS method. Note that FEA simulations are performed using 8 CPUs while only one CPU is used for the LVLS method. Even though the LVLS method performs stress analysis on a 2D plane (= device layer), whereas FEA simulation is performed on the entire 3D structure, stress analysis for other planes can be performed in a similar way if needed.

**Table 28:** Von Mises stress comparison between FEA and LVLS for a 4-die stack structure (die0). Error = LVLS - FEA. (At TSV edge, typical von Mises stress level is around 900MPa.)

# TSV / $\mu$ -B /pkg-B	FEA		LVLS		max error (MPa)		
	# node	run time	# grid	run time	inside TSV	TSV edge	outside TSV
1/1/1	754K	1d2h	1M	23s	-11.4	-12.6	7.9
2/2/1	812K	1d2h	1M	26s	-12.7	-13.2	7.3
5/5/2	902K	1d6h	6M	2m43s	-14.1	-15.3	8.2
10/10/4	1.3M	1d20h	9M	6m44s	-23.1	-19.8	9.4
10/10/9	1.4M	2d0h	16.8M	11m11s	-22.5	-20.5	11.9

The error between FEA simulations and LVLS is very small. Results show that our LVLS method underestimates stress magnitude inside TSV and TSV edge and overestimates outside TSV. In general, the most critical region for the mechanical reliability is the interface between different materials, hence TSV edge is most important in our case. Even though the maximum error at TSV edge is as high as -20.5MPa, its % error is only -2.24%. Figure 48 shows one test case comparison of von Mises stress between FEA and LVLS. The structure has 10 TSVs ( $5\mu m$  diameter and  $10\mu m$  pitch),  $10\mu$ -bumps ( $20\mu m$  diameter and  $40\mu m$

pitch), and 9 package-bumps (100 $\mu m$  diameter and 200 $\mu m$  pitch). It clearly shows the LVLS method matches well with the FEA simulation result.



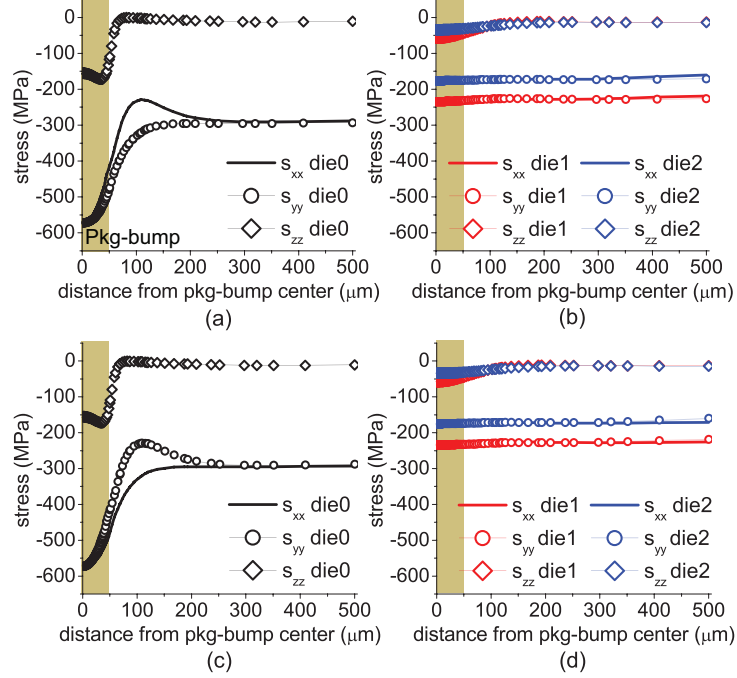
**Figure 48:** Sample stress comparison between FEA and LVLS. (a) Test structure. (b) Close-up shot of von Mises stress map (using LVLS) taken from the red box in (a) on the device layer in die0 in a 4-die stack. (c) FEA vs. LVLS along the red line in (b).

## 5.5 Full-Stack Timing Variation Analysis

### 5.5.1 Full-Stack Device Mobility Variation

From FEA simulations, a highly compressive stress is observed on device layers due to package-bumps, which is induced by the CTE mismatch between package-bumps and underfill. As shown in Figure 49, die0 (= closest to package-bump layer) experiences the most compressive stress due to their proximity. The stress becomes less compressive as we go to upper dies. The stress distribution ( $\sigma_{xx}$  and  $\sigma_{yy}$ ) in die3 (un-thinned top die) is almost flat ( $-110 \pm 5$ MPa), since die3 does not contain any TSVs, which is discussed in Section 5.2.2. Thus, only the stress in the dies containing TSVs is computed.

In Equation (17), the electron mobility variation is approximately proportional to the sum of  $\sigma_{xx}$  and  $\sigma_{yy}$  due to the same sign (= negative) of  $\pi'_{11}$  and  $\pi'_{12}$ , while the hole mobility variation is roughly proportional to the difference between  $\sigma_{xx}$  and  $\sigma_{yy}$  due to the opposite sign of  $\pi'_{11}$  and  $\pi'_{12}$ . The stress distribution on device layers induced by package-bump only are shown in Figure 49. Although there is a noticeable difference between  $\sigma_{xx}$  and  $\sigma_{yy}$  near the package-bump edge in die0, their difference is almost negligible in other regions. Thus, this package-bump induced stress will not alter the hole mobility variation significantly except near the package-bump edge area. On the other hand, the electron mobility will be degraded under the influence of the package-bump since both  $\sigma_{xx}$  and  $\sigma_{yy}$  are compressive

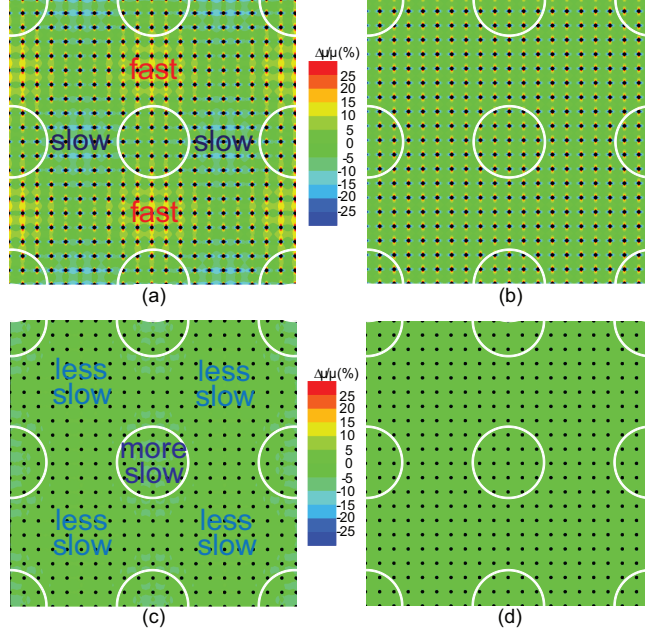


**Figure 49:** Normal stress components induced by package-bump on device layers (FEA results). (a) Stress in die0 along the x-direction. (b) Stress in die1 and die2 along the x-direction. (c) Stress in die0 along the y-direction. (d) Stress in die1 and die2 along the y-direction.

(= negative), which is shown in Figure 45(d). Furthermore, the level of electron mobility degradation is most severe in die0.

Figure 50 shows hole and electron mobility variation maps in a four-die stack with 441 TSVs/ $\mu$ -bumps with  $20\mu m$  pitch and 9 package-bumps with  $200\mu m$  pitch. Both hole and electron mobility variation range is largest in die0 due to the direct impact of package-bump-induced stress. Especially, the hole mobility degrades in between package-bumps along the x-direction and improves along the y-direction. This is because of the difference between  $\sigma_{xx}$  and  $\sigma_{yy}$  stress components near package-bump edge area shown in Figure 49(a): along the x-direction  $\sigma_{xx}$  is higher than  $\sigma_{yy}$ , while along the y-direction  $\sigma_{yy}$  is higher than  $\sigma_{xx}$ . The electron mobility in die0 degrades in most cases, and the worst spot is inside the package-bump area since the most compressive stress occurs in this region as shown in Figure 49.

In addition, Figure 50 shows that the stress induced by package-bumps affects the mobility variation of a large number of cells, while TSVs generate the mobility variation



**Figure 50:** Mobility variation map with 441 TSVs/ $\mu$ -bumps (black dots) and 9 C4 bumps (white circles) (LVLS results). (a) Hole mobility variation map in die0. (b) Hole mobility variation map in die2. (c) Electron mobility variation map in die0. (d) Electron mobility variation map in die2.

pattern only for the cells nearby these TSVs. It is also observed that as we go to upper dies, mobility variations due to package-bumps are almost negligible, hence the mobility variation pattern is mostly determined by TSVs.

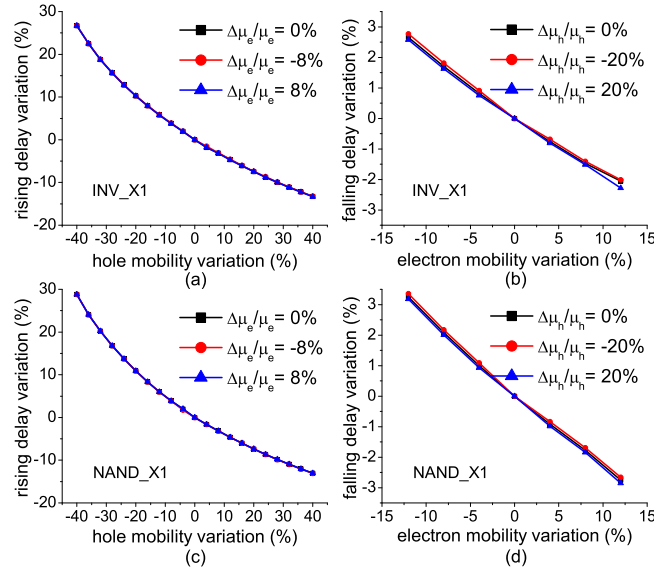
### 5.5.2 Chip/Package Stress-Aware Timing Analysis

In this section, the stress-aware static timing analysis (STA) flow is presented. First, a Verilog netlist and a parasitic extraction file (SPEF) for each die are built from 3D IC layouts. Each instance name in the netlists is replaced by corresponding hole and electron mobility variation based on our stress and mobility analysis results. For example, INV\_X1 with +4% hole mobility and -8% electron mobility variation becomes INV\_X1\_Hp4\_Em8. Then, a top-level Verilog netlist that instantiates each die design and connects the 3D nets using TSV is created. In addition, a top-level SPEF file that contains parasitic models of the TSVs is generated. Lastly, 3D STA is performed using Synopsys PrimeTime [56].

For this stress-aware STA, a timing library is built to capture the mobility variation impact on cell delay. First, both hole and electron mobility variation range affected by

multiple TSVs,  $\mu$ -bumps, and package-bumps is obtained. Since this range is different across the stack and also affected by the alignment and the pitch of TSVs,  $\mu$ -bumps, and package-bumps, several test cases are generated by varying these knobs. One of the test cases is shown in Figure 50. The hole mobility varies from -52% to 52% and the electron mobility ranges from -16% to 8% without any TSV keep-out-zone (KOZ), where devices cannot be placed. Actual mobility variation range is reduced by introducing KOZ. Cell timing with the mobility variation is characterized using Cadence Encounter Library Characterizer [57] with a 2% mobility step size.

Figure 51 shows the FO4 delay of INV\_X1 and NAND\_X1 gates with mobility variations. It is observed that the delay variation range is similar for both gates with given mobility variations. Note that the rise delay is not affected by the electron mobility variation and the fall delay is not much influenced by the hole mobility change. Thus,  $\Delta\mu_e/\mu_e$  can be fixed when  $\Delta\mu_h/\mu_h$  is swept, and vice versa. This is useful to reduce the number of library characterization. Instead of characterizing 689 ( $= 53 \times 13$  with 2% step size) libraries, 66 ( $= 53 + 13$ ) libraries need to be prepared [11].



**Figure 51:** Mobility variation impact on cell FO4 delay. (a) Rise delay dependency on hole mobility variation (INV\_X1). (b) Fall delay dependency on electron mobility variation (INV\_X1). (c) Rise delay dependency on hole mobility variation (NAND\_X1). (d) Fall delay dependency on electron mobility variation (NAND\_X1).



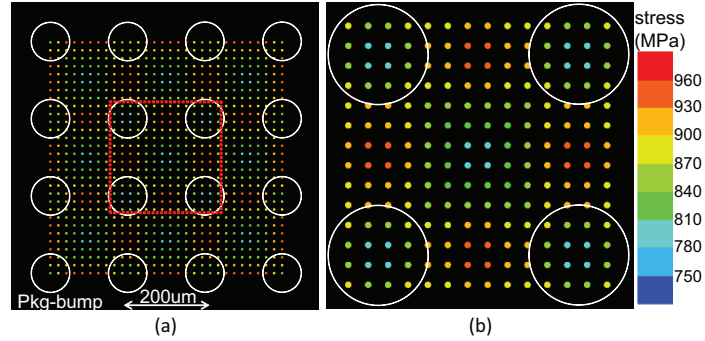
## 5.6 Full-Stack Reliability Analysis Results

A chip/package thermo-mechanical stress and reliability co-analysis flow based on LVLS is built in C++/STL. The impact of package-bump and  $\mu$ -bump on the reliability in full-system scale is explored. In addition, the reliability concerns in wide-I/O DRAM and block-level 3D IC designs are examined.

In the experiments, a regular TSV placement style in which TSVs are placed uniformly across each die or inside TSV blocks with pre-defined pitch is adopted. In all cases, the pair of TSV and  $\mu$ -bump is vertically aligned. Default diameter/height ( $\mu\text{m}$ ) of TSV,  $\mu$ -bump, and package-bump are 5/30, 10/10, and 100/100, respectively, unless otherwise specified.

### 5.6.1 Impact of Package-Bump and $\mu$ -Bump

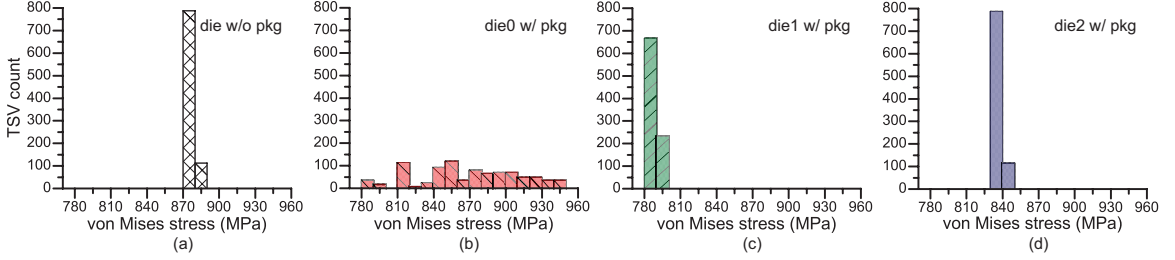
The impact of package-bump and  $\mu$ -bump on the mechanical reliability of different dies in a four-die stack is studied. This is also compared to the case without these components as in the previous work [5] as shown in Figure 38(a). In this experiment, the pitch of TSV/ $\mu$ -bump and package-bump are  $20\mu\text{m}$  and  $100\mu\text{m}$ , respectively; the total number of TSV/ $\mu$ -bump and package-bump are 900 and 16, respectively, as shown in Figure 52.



**Figure 52:** Von Mises stress map for TSVs (die0 in a 4-die stack). Colored dots are TSVs and white circles are package-bumps. (a) Test structure. (b) Close-up shot of red box in (a)

It is first observed that unlike the die without package-bumps and  $\mu$ -bumps (Figure 53(a)) and the upper dies with package components (Figure 53(c) and (d)), TSVs in die0 (Figure 53(b)) experience large variations of von Mises stress across the die. This is because die0 is highly affected by package-bumps underneath it, and hence depending on

the relative position between TSVs in die0 and package-bumps the von Mises stresses of TSVs change noticeably. Note that higher von Mises stress level is observed in (Figure 53(a) than the previous work [5] even with the same simulation structure. This is because the Young's modulus of 188GPa for Si is used instead of 130GPa in [5] as a worst case scenario.



**Figure 53:** Impact of package components and die stacking on the mechanical reliability of TSVs (900 TSVs in each die).

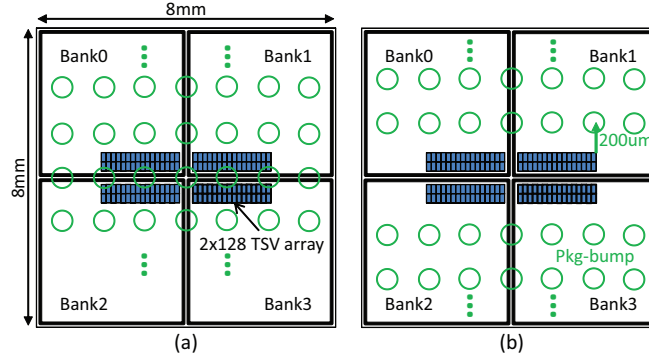
It is also identified that higher von Mises stress occurs around package-bump edge and in between package-bumps due to constructive stress interference shown in Figure 52(b). However, as shown in the center of Figure 52(b), if the distance between TSV and package-bumps is long enough, the von Mises stress of TSV becomes low.

Interestingly, die1 shows the lowest von Mises stress level among all cases even though die2 is farthest from package-bumps. This is due to the fact that die2 is affected by the rigid un-thinned top silicon substrate above it. Since die0 is most problematic in terms of the mechanical reliability, we only consider die0 in a four-die stack in the subsequent experiments.

### 5.6.2 Case Study I: Wide-I/O DRAM

Wide-I/O based 3D DRAM is fast becoming the first mainstream product that utilizes TSV in 3D ICs, mainly targeting mobile computing applications such as smart phones which need lower power consumption and high data bandwidth. In this section, the reliability concerns of TSVs in wide-I/O DRAM are evaluated. The TSV placement style similar to the work in [58] is followed, where TSV arrays are placed in the middle of a chip. It is assumed that 2x128 TSV array (per memory bank) is placed in the middle of a chip shown in Figure 54. Four memory banks and 1024 TSVs in total are employed. The pitch of TSV/ $\mu$ -bump

and package-bump are set as  $15\mu m$  and  $200\mu m$ , respectively. Two cases are compared; (a) Package-bumps are placed right underneath TSV arrays; (b) Package-bumps are placed with  $200\mu m$  spacing from TSV arrays. This  $200\mu m$  distance is chosen since the effect of package-bump on the TSV reliability is negligible beyond  $200\mu m$  in case of the  $100\mu m$  diameter package-bump as shown in Figure 43.



**Figure 54:** Mechanical reliability in wide I/O DRAM. 1024 TSVs are placed in the middle of a chip. (a) Package-bumps are placed underneath TSV arrays. (b) Package-bumps are placed  $200\mu m$  apart from TSV arrays. (not drawn in scale.)

Table 29 clearly shows that the chip/package co-design can greatly reduce the mechanical reliability concerns in TSV-based 3D ICs. With a safe margin of  $200\mu m$  (= case(b)), von Mises stress magnitude reduces significantly. Thus, given the TSV placement, we can find safe locations for package-bumps without affecting the package design much, or vice versa.

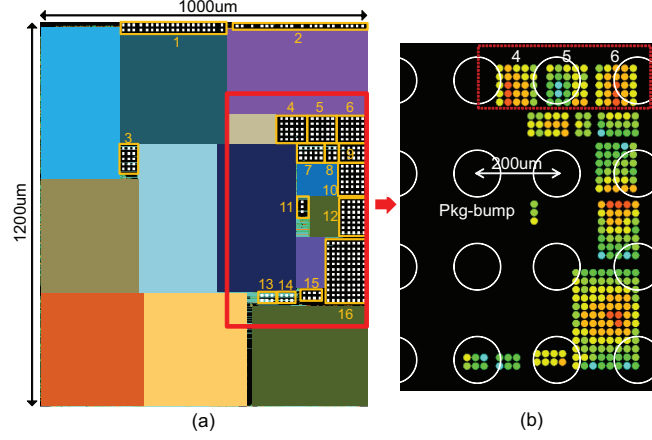
**Table 29:** Reliability in wide-I/O DRAM.

case	von Mises stress distribution (MPa)					median (MPa)
	780-810	810-840	840-870	870-900	900-930	
(a)	30	114	52	220	608	944.8
(b)	182	842	0	0	0	856.2

### 5.6.3 Case Study II: Block-Level 3D IC

Now the reliability issues in block-level 3D designs are examined. 3D block-level designs are generated using an in-house 3D floorplanner which treats a group of TSVs as a block shown in Figure 55. Total 16 TSV blocks (368 TSVs) are used and the TSV pitch is  $15\mu m$ . Package-bumps are regularly placed with  $200\mu m$  pitch.

Table 30 shows von Mises stress level in selected TSV blocks. It is first observed that



**Figure 55:** Mechanical reliability in block-level 3D IC. (a) Sample layout of block-level design. (b) Von Mises stress map for TSVs in red box in (a).

larger TSV blocks experience more variation of von Mises stress within the TSV block. This is because the distance between each TSV in the block and package-bumps can vary more than small TSV blocks, which is a key factor that affects the reliability of TSVs. In addition, TSV blocks with the same size show quite different characteristics depending on the distance to the nearest package-bump. For example, although TSV block 4, 5, and 6 are all 5x5 TSV blocks and are located side-by-side, TSV block 5 shows the lowest von Mises stress level. However, its standard deviation of von Mises stress is highest among three blocks. Lower von Mises stress is observed if TSVs are placed near the package-bump center or far away from it; however, higher stress is observed in TSVs located around package-bump edge as shown in Figure 43. In case of TSV block 5, most TSVs are near the package-bump center, which lowers von Mises stress level. However, at the same time a few TSVs are around the package-bump edge, which increases the standard deviation of von Mises stress inside the TSV block.

**Table 30:** Mechanical reliability in block-level 3D IC. TSV blocks are shown in Figure 55.

TSV block #	# TSV	von Mises stress (MPa)				blk-bump dist ( $\mu m$ )
		max	min	avg	std dev	
3	5x3	901.0	811.1	859.5	26.0	96.4
4	5x5	939.6	853.5	902.6	24.0	67.6
5	5x5	908.6	816.0	858.7	33.3	24.1
6	5x5	942.3	874.4	910.4	22.0	91.4
11	3x1	896.6	855.9	871.0	18.2	39.3
16	12x8	943.7	806.0	877.2	33.6	90.7

From this experiment, two possible ways are identified to reduce the mechanical reliability problems in block-level 3D designs: (1) Assign TSV blocks right above package-bump center locations if possible. (2) Place package-bumps outside the TSV block locations with a safe margin such as outside the red box in Figure 55(a). However, other design constraints such as package area and the required number of pins should be carefully considered as well.

## 5.7 Full-Stack Timing Analysis Results

In this section, the impact of chip/package elements on the full-stack timing results is investigated. In the simulations, four-die stack 3D IC designs are designed using Cadence Encounter [23] with Nangate 45nm cell library [32]. A regular TSV placement style is employed.

In all cases, a pair of TSV and  $\mu$ -bump is always vertically aligned. The default diameter/height ( $\mu m$ ) of TSV,  $\mu$ -bump, and package-bump are 5/30, 10/10, and 100/100, respectively, unless otherwise specified. The package-bump pitch is assumed to be 200 $\mu m$  for all cases.

### 5.7.1 2D vs. 3D Stress Impact on Mobility and Timing

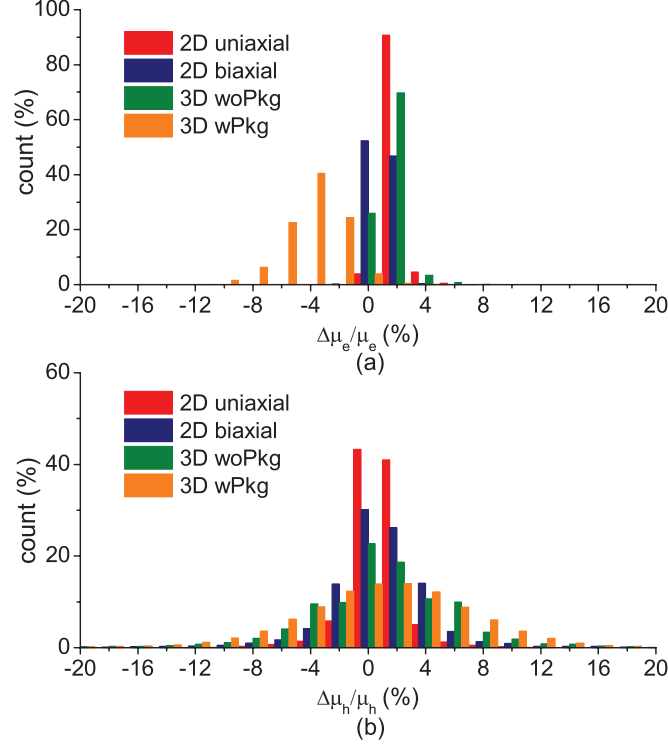
First, the impact of different stress cases, i.e., 2D stress (*2D uniaxial* and *2D biaxial*) and 3D stress (*3D woPkg* and *3D wPkg*), on the full-stack timing and mobility variations is examined. Three circuits listed in Table 31 with the TSV KOZ size of 1 $\mu m$  are employed. Note that all benchmark circuits are designed with the timing optimization objective, but the stress impact is not considered in design stages.

**Table 31:** Benchmark circuits with TSV KOZ = 1.0 $\mu m$

circuit	# cell	area ( $\mu m \times \mu m$ )	WL (mm)	# TSV	TSV pitch ( $\mu m$ )	profile
ckt1	51K	$290 \times 290$	1235	1062	15	DES
ckt2	592K	$800 \times 800$	15831	2325	20	512pt FFT
ckt3	1.31M	$1150 \times 1150$	36842	6632	25	1024pt FFT

Figure 56 shows the cell mobility distribution in die0 in a four-die stack of ckt2. It is observed that the electron mobility is highly concentrated within 0~2% range for both 2D stress cases and the *3D woPkg*. Note that the *2D uniaxial* case always improves the

electron mobility, while the *2D biaxial* case can degrade the electron mobility as well. Most importantly, the electron mobility variation with package components shows quite a different behavior: the mobility variation range is wider than other cases and most of cells in die0 experience the electron mobility degradation. The degradation is mainly due to the compressive stress from package-bumps. Also, the wider distribution originates from the relative positions between cells, TSVs/ $\mu$ -bumps, and package-bumps.

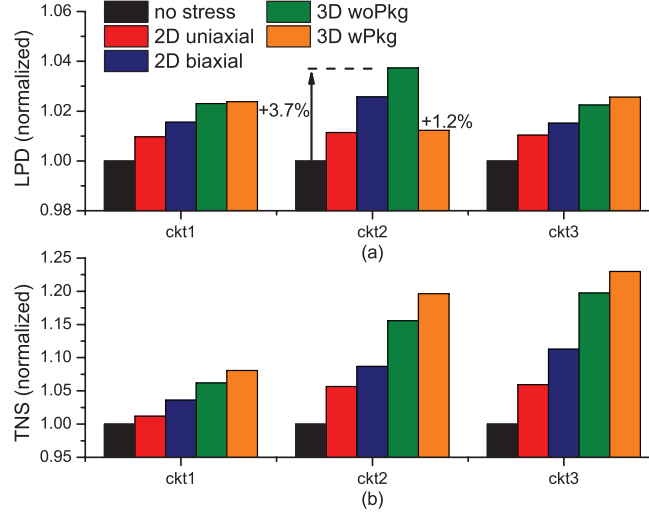


**Figure 56:** Cell mobility variation histogram in die0 in 4-die stack (ckt2). (a) Electron mobility. (b) Hole mobility.

As for the hole mobility distribution, all cases show wider distribution than the electron mobility case, which is expected from Figure 45. However, still the *3D wPkg* case generates the largest variation, which is clear as shown in Figure 50(a). Note that as we go to upper dies, the hole mobility distribution of the *3D wPkg* becomes comparable to the *3D woPkg* case.

Stress-aware 3D STA results are shown in Figure 57. The longest path delay (LPD) and total negative slack (TNS) for different stress cases are monitored. First, it is observed that the *2D uniaxial* case always underestimates the LPD compared with the *2D biaxial*

case. Interestingly, the LPD of ckt2 in the *3D wPkg* case shows better timing than other stress cases shown in Figure 57(a). This can be explained with Figure 58. As Figure 58(a) shows, the cells in the critical path are located in between package-bumps in y-direction. In this case, the hole mobility improves as shown in Figure 50(a). Moreover, the hole mobility further improves when cells are placed in between TSVs in y-direction as shown in Figure 58(d).

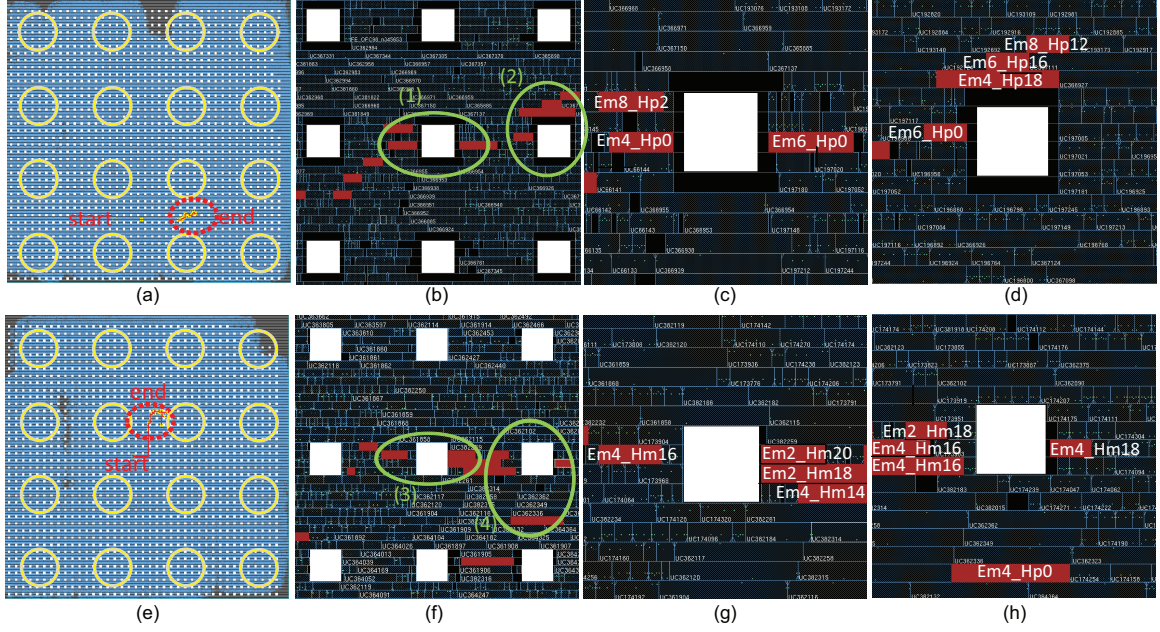


**Figure 57:** Impact of 2D and 3D stress cases on the longest path delay (LPD) and total negative slack (TNS). Timing numbers are normalized to the no-stress case. TSV KOZ is  $1 \mu m$  for all cases. (a) LPD variation. (b) TNS variation.

The opposite case can also happen as shown in Figure 58(e), where the cells in the critical path are placed in between package-bumps in x-direction. In this case, the LPD degrades by 5.2% in the *3D wPkg* case compared with the *no-stress* case, while the *3D woPkg* case degrades the LPD by 2.3%.

The impact of package-bump stress on the mobility in die0 is clear if we compare Figure 58(c) and 58(g). Although the relative positions between TSV and cells are similar, the hole mobility variation is significantly different depending on package-bump locations.

The stress impact on timing is more evident in TNS. In the *3D wPkg* case, TNS is larger than the *no-stress* case up to 22.9% as shown in Figure 57(b). This is because most cells in the design are affected by the stress induced by TSVs,  $\mu$ -bumps, and package-bumps, and



**Figure 58:** Full-chip layout (die0 in 4-die stack) with the highlighted longest path. White squares are TSVs and yellow circles are package bumps. (cell mobility naming convention: e.g., Em8.Hp4 = electron mobility minus 8 % and hole mobility plus 4 %). (a) Layout of ckt2 (KOZ =  $1.0\mu m$ ). (b) Cells in red circle in (a). (c) Close-up shot of green circle (1) in (b). (d) Close-up shot of green circle (2) in (b). (e) Layout of ckt2 (KOZ =  $0.3\mu m$ ). (f) Cells in red circle in (e). (g) Close-up shot of green circle (3) in (f). (d) Close-up shot of green circle (4) in (f).

thus undergo mobility variations.

### 5.7.2 Case Studies: Block-level and Wide-I/O Style 3D Designs

In this section, the chip/package stress impact on the full-stack timing in block-level and wide-I/O style designs listed in Table 32 is studied. In case of the block-level design, the high mobility variation region is limited to nearby TSV blocks. Although the global mobility variation pattern is largely determined by package-bumps, the local mobility minima and maxima are mostly caused by TSVs. Thus, most of the cells inside functional blocks do not experience high mobility variations.

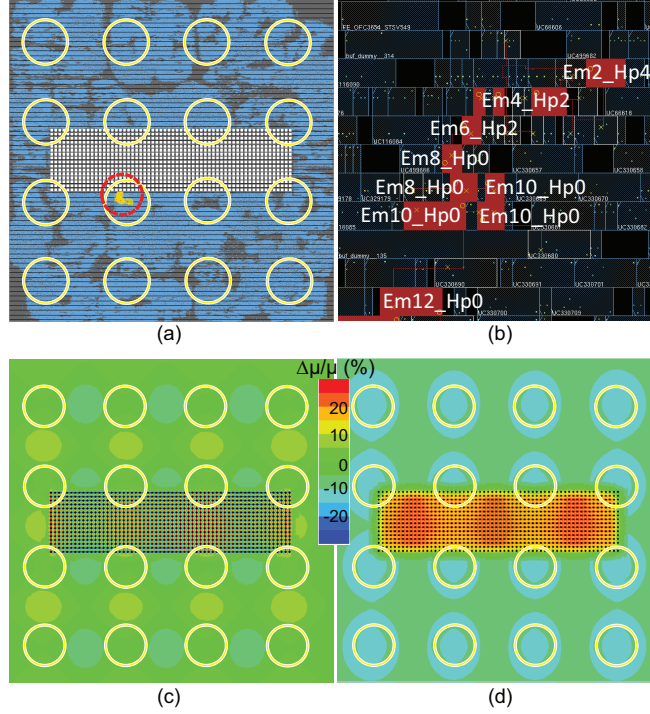
In case of the wide-I/O style design, it is assumed that 8x30 TSV array (per memory bank) is placed in the middle of a chip. In addition, there are four memory banks, hence the total 960 TSVs are employed in die0 as shown in Figure 59. The hole and electron mobility maps in Figure 59(c) and 59(d) clearly show that high mobility variation region



is confined to inside and nearby the TSV array. Thus, majority of cells are not affected by the TSV stress similar to the block-level design.

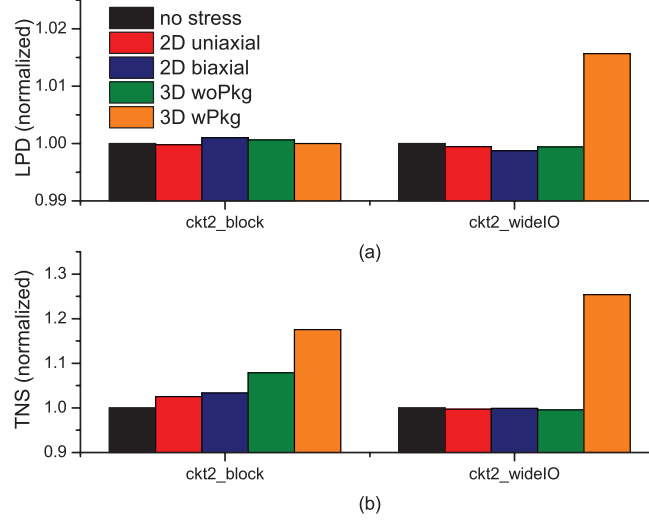
**Table 32:** Block-level and wide-I/O style 3D IC designs. Package-bump pitch is  $200\mu m$ .

circuit	# cell	area ( $\mu m \times \mu m$ )	WL (mm)	# TSV	TSV pitch ( $\mu m$ )
ckt2_block	578K	$840 \times 920$	16083	1769	15
ckt2_wideIO	578K	$820 \times 820$	15521	2116	10



**Figure 59:** Layout and mobility variation map of wide-I/O style design (ckt2-wide). (a) Layout of die0 in 4-die stack with the highlighted cells in the critical path. (b) Close-up shot of red circle in (a). (c) Hole mobility variation map. (d) Electron mobility variation map (LVLS results).

Figure 60 shows 3D STA results for the block-level and the wide-I/O style designs. As for the LPD, there is an almost negligible impact from all stress cases for both block-level and wide-I/O style designs, since most cells are not affected by the TSV stress. One exception is the *3D wPkg* case in the wide-I/O style design. This is because the cells in the critical path are placed nearby TSV array and right above a package-bump as shown in Figure 59(a). Cells that are placed in vertical direction with respect to TSVs experience



**Figure 60:** Impact of 2D and 3D stress cases on the longest path delay (LPD) and total negative slack (TNS) in block-level and wide-I/O style 3D IC designs. TSV KOZ is  $1.7\mu m$  for all cases. (a) LPD variation. (b) TNS variation.

electron mobility degradation and hole mobility improvement. However, the electron mobility further decreases inside package-bump area as shown in Figure 50(c), hence the net effect is timing degradation.

In addition, more TNS variation is observed in the block-level design than that in the wide-I/O style design for *2D uniaxial*, *2D biaxial*, and *3D woPkg* cases. The block-level design contains more TSV blocks than the wide-I/O style design, hence the number of cells nearby these TSV blocks also increases. Thus, more paths are affected by the TSV stress than the wide-I/O style design. However, as the impact of package-bumps is included, all cells in these designs are affected by package-bumps, hence non-negligible variations in TNS are observed for both design styles.

## 5.8 Summary

In this chapter, it is demonstrated that how package elements affect the stress field and the mechanical reliability on top of the TSV-induced stress in 3D ICs. In addition, the impacts of chip and package components on the mobility and full-stack timing variations in 3D ICs are studied. It is observed that the mechanical reliability of TSVs in the bottom-most die in the stack are highly affected by packaging elements, and that effect decreases

as we go to the upper dies. Moreover, an accurate and fast full-chip/package stress and mechanical reliability co-analysis flow is presented based on the principle of lateral and vertical linear superposition of stress tensors (LVLS), considering all chip/package elements. Lastly, a chip/package stress-aware timing analysis method is presented, which is applicable to stress-aware full-stack timing optimization for 3D ICs.

## CHAPTER VI

### CRACK ANALYSIS AND OPTIMIZATION FOR THROUGH-SILICON-VIA-BASED 3D ICS

There have been major concerns on the thermo-mechanical reliability of TSV structures. If there is a small defect such as a void around a TSV, the TSV-induced stress can drive the interfacial cracking between dielectric liner and silicon substrate or the cohesive cracking in dielectric liner and silicon substrate. These cracks may damage transistors nearby, create conducting paths between TSVs (= short circuit), and cause the entire chip operation failure in the worst case. Previous works studied the crack growth behavior under TSV stress [12, 13]. However, most previous works focused on modeling the thermo-mechanical stress and reliability of a single TSV in isolation. These simulations were performed using FEA methods which are computationally expensive or infeasible for full-chip-scale analysis.

The work in [33] investigated cohesive cracks in Si substrate with straight and zigzag TSV lines containing five TSVs, and showed that the zigzag type is a better choice to mitigate crack driving force than the straight line. However, this work was performed based on a 2D stress model. Thus, it does not capture the 3D nature of a stress field near a wafer surface around TSVs where devices are located [13]. Moreover, they did not consider a dielectric liner and a landing pad in TSV structures, which are essential components for TSVs.

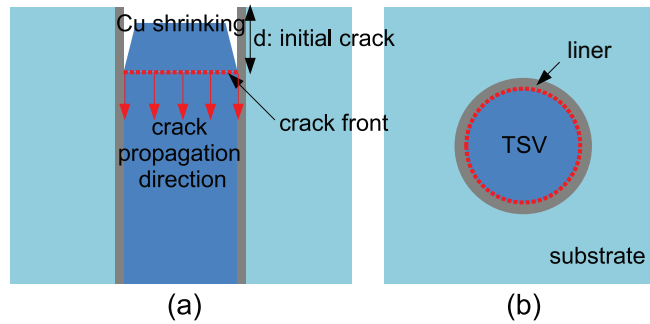
Although authors in [13] proposed a semi-analytic ERR model for TSV interfacial crack, it is only valid for an infinitely long TSV. Moreover, their model is only applicable to a single TSV in isolation and their TSV structure includes TSV and silicon substrate only. Thus, their model cannot be directly used to assess TSV interfacial cracks considering multiple TSVs as well as the TSV which contains a landing pad and a dielectric liner because of the change in boundary conditions.

In this chapter, a fast and efficient full-chip TSV interfacial crack analysis flow is presented based on design of experiments (DOE) and response surface methodology (RSM). An energy release rate (ERR) is used as a mechanical reliability metric and the impact of TSV placement style on the ERR is examined. In addition, detailed studies on the substrate crack around TSV are provided. The impact of TSV placement structures on the substrate crack growth patterns is also discussed.

### 6.1 TSV Interfacial Crack Modeling

It is widely known that most of the mechanical reliability failures occur at the interface between different materials. Thus, in this section, the TSV interfacial crack at the TSV/dielectric liner interface is investigated. However, it is hard to obtain realistic crack structures and crack growth behavior models after crack initiation without measurement data. Even with the same initial crack, every crack can grow in a different manner depending on the surrounding environment.

Therefore, a crack structure well studied from previous works [13, 14] is adopted. The TSV interfacial crack structure is shown in Figure 61. This crack initiates around the circumference of the TSV near the wafer surface and grows vertically downward. In this study, this wafer surface means the dielectric layer surface right below the dielectric layer ( $\text{SiO}_2$ )/ILD (low K) interface. In addition, it is assumed that the crack front propagates uniformly to simplify crack modeling.



**Figure 61:** TSV interfacial crack structure under a negative thermal load. (a) Side view with initial crack length of  $d$ . (b) Top view.

### 6.1.1 Energy Release Rate

The energy release rate (ERR) is defined as the energy dissipated during fracture, i.e. crack, per newly created fracture surface area. According to the law of conservation of energy, energy supplied to a crack front for it to grow must be balanced by the amount of energy dissipated due to the formation of new surfaces. In other words, ERR is the measure of the amount of energy available for fracture. If high energy is available around crack front, then there is a high chance of crack growth.

However, even with an initial crack, if the ERR of the crack under consideration is lower than a threshold value, this crack does not grow further and stays in a stable state. The debonding energy between different materials is this threshold, and its value is material and fabrication process specific. For example, the debonding energy of Cu/SiO<sub>2</sub> interface ranges from 0.7 to 10 J/m<sup>2</sup> depending on the fabrication process [12].

Since the loading in this simulation structure is solely due to the thermal expansion from the fabrication process with no work done by external loads, the ERR can be determined as the rate of change in strain energy with crack extension [12]. In TSV-based 3D ICs, this strain energy is mostly generated from the thermo-mechanical stress induced by TSVs. Based on this, two 3D FEA models are created for strain energy analysis, one with a crack length of  $d$ , and another with a crack length of  $d + \Delta d$ . The ERR for TSV interfacial crack is obtained using a forward difference approach as follows:

$$ERR = -\frac{\partial U}{\partial A} = -\frac{U_{d+\Delta d} - U_d}{2\pi r_{TSV} \cdot \Delta d}$$

where,  $U$  is a strain energy,  $A$  is area,  $d$  is an initial crack length,  $\Delta d$  is a crack increment, and  $r_{TSV}$  is TSV radius. In this study,  $d$  is set as 1  $\mu m$  and  $\Delta d$  is varied from 0.1 to 0.5  $\mu m$  to capture the crack initiation from the wafer surface as well as near surface thermal stress impact [13].

Before discussing the detailed crack modeling, two terminologies are introduced: (1) **Victim TSV**: TSV with an interfacial crack. (2) **Aggressor TSV**: TSV located nearby a victim TSV and affecting crack growth of the victim TSV.

### 6.1.2 3D FEA Simulation

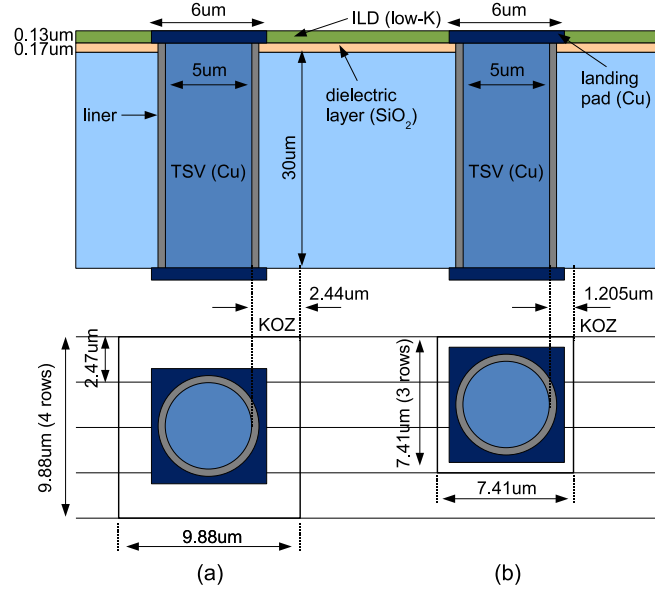
Since there is no known analytical ERR model for a realistic TSV structure, 3D FEA models for a TSV interfacial crack are created to investigate the impact of aggressor TSVs on the interfacial crack of a victim TSV. To realistically examine the interfacial crack, the baseline simulation structure of a TSV is based on the fabricated and the published data [24], as shown in Figure 62.

Two TSV cells are constructed, i.e., TSV<sub>A</sub> and TSV<sub>B</sub>, which occupy four and three standard cell rows in 45nm technology. The We keep-out-zone (KOZ) in which no cell is allowed to be placed is defined as 2.44 $\mu m$  and 1.205 $\mu m$  from TSV edge for TSV<sub>A</sub> and TSV<sub>B</sub> cells, respectively. The baseline TSV diameter, height, landing pad width, and liner thickness are 5 $\mu m$ , 30 $\mu m$ , 6 $\mu m$ , and 125nm, respectively, unless otherwise specified, which are close to the data in [24]. SiO<sub>2</sub> is used as a baseline liner material, and the Cu diffusion barrier material such as Ta and Ti is ignored in these experiments.

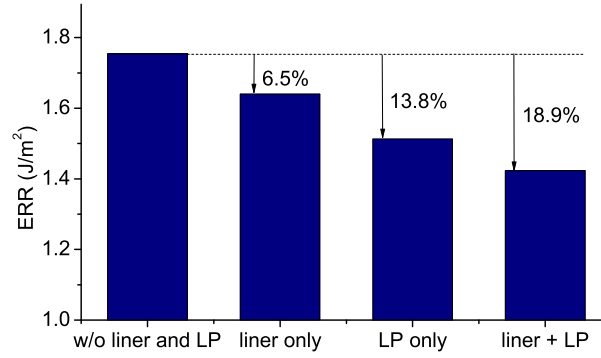
Material properties used for the simulations are as follows: CTE (ppm/K) for Cu = 17, Si = 2.3, SiO<sub>2</sub> = 0.5, and BCB = 40; Young's modulus (GPa) for Cu = 110, Si = 130, SiO<sub>2</sub> = 71, and BCB = 3. A FEA simulation tool ABAQUS [37] is used to perform simulations.  $\Delta T = -250^\circ C$  of thermal load is applied for the entire simulation structure. That is, the TSV structure is annealed at 275 $^\circ C$  and cooled down to 25 $^\circ C$  to represent the manufacturing process [13,33,39]. It is also assumed that the entire TSV structure is stress free at the annealing temperature.

### 6.1.3 Impact of TSV Liner and Landing Pad

In this section, the impact of surrounding structures such as a liner and a landing pad on the crack growth is explored. For this experiment, the simulation structures without aggressor TSVs are employed. Figure 63 shows the ERR of a victim TSV with four different configurations. First, the 125nm thick SiO<sub>2</sub> liner, which acts as a stress buffer layer, reduces the ERR by 6.5% compared with the case without a landing pad and a liner. The landing pad also helps decrease the ERR by preventing TSV/liner interface from separating. Finally, when both a SiO<sub>2</sub> liner and a landing pad are considered, the ERR decreases by 18.9%.



**Figure 62:** Baseline TSV structure. (a) TSV<sub>A</sub> cell occupying four standard cell rows (KOZ = 2.44μm). (b) TSV<sub>B</sub> cell (KOZ = 1.205μm).

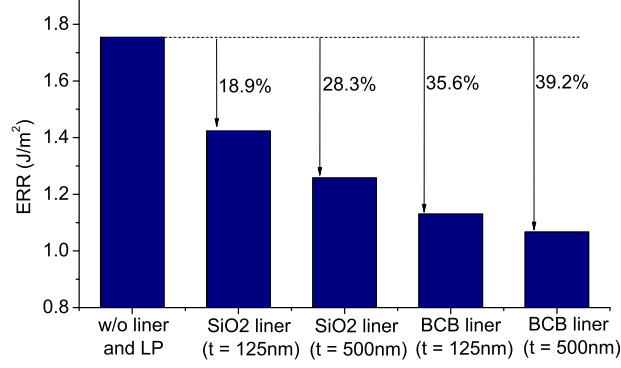


**Figure 63:** Impact of TSV surrounding structures on ERR.

In addition, benzocyclobutene (BCB), a polymer dielectric material, is employed as an alternative TSV liner material [13, 33]. Since Young's modulus, which is a measure of the stiffness of an isotropic elastic material, of BCB is much lower than Cu, Si, and SiO<sub>2</sub>, this BCB liner can absorb the stress effectively caused by the CTE mismatch. The impact of the liner material and its thickness on ERR is shown in Figure 64. For this experiment, the landing pad width is set as 6μm for all cases. As the liner thickness increases, the ERR decreases noticeably for both liner materials. More importantly, the BCB liner outperforms SiO<sub>2</sub> on reducing the ERR.

The impact of landing pad size on the TSV interfacial crack is examined as well. Four





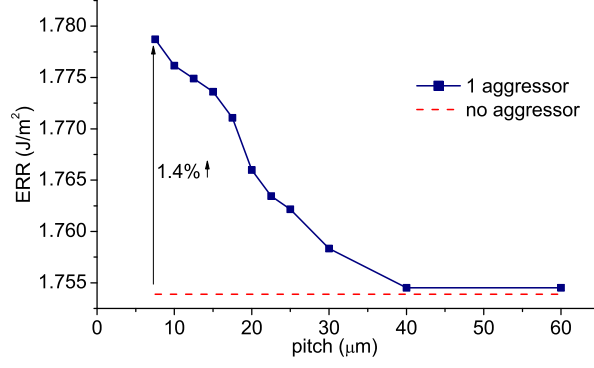
**Figure 64:** Impact of TSV liner on ERR. Landing pad width of  $6\mu m$  is used with each liner case.

landing pad widths are used; 6, 8, 10, and  $12\mu m$ . The ERR with a landing pad is lower for all landing pad sizes than the ERR without a landing pad case. However, it is also observed that for the landing pad width up to  $10\mu m$ , the ERR increases and then saturates. This is because the magnitude of all normal stress components at TSV/liner interface underneath the landing pad increases due to the increased Cu volume with enlarged landing pad area, which is an additional CTE mismatch source. Thus, for the TSV interfacial crack, the larger landing pad size is not beneficial.

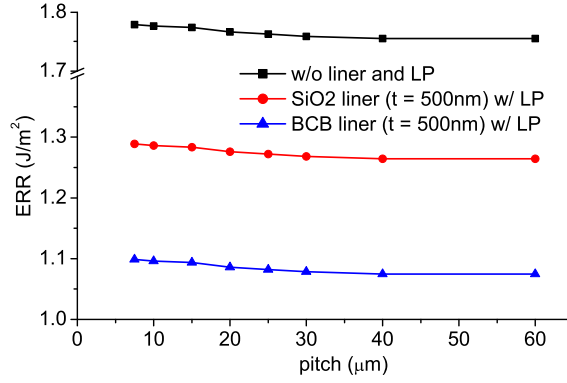
#### 6.1.4 Impact of Pitch and Angle among TSVs

First, the impact of TSV pitch on the TSV interfacial crack is investigated in this section. With a fixed victim TSV location, the TSV pitch between a victim and an aggressor is varied from  $7.5\mu m$  to  $60\mu m$ . As shown in Figure 65, the ERR decreases monotonically as the pitch increases and approaches to the level when there is no aggressor at around  $40\mu m$  pitch. However, when only one aggressor is considered, the increase of ERR at the minimum pitch compared with the maximum pitch is only 1.4%, which is negligible. The ERR curves with two  $500nm$  thick liner materials and a  $6\mu m$  wide landing pad are shown in Figure 66. The magnitude of ERR decreases when a liner and a landing pad are considered, but the overall ERR trend remains similar.

As additional TSVs are introduced, both distance and angle between TSVs become important to the TSV interfacial crack. The stress at a point can be computed by adding individual stress tensors induced by each TSV at this point. Depending on a relative angle



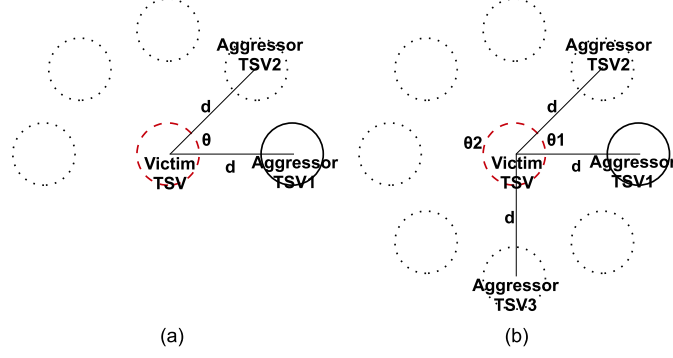
**Figure 65:** Impact of TSV pitch on ERR.



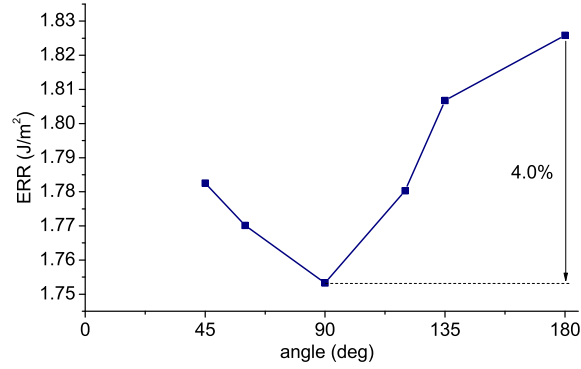
**Figure 66:** Impact of TSV pitch with liner and landing pad ( $6\mu m$ ) on ERR.

between TSVs, even with the same pitch, the stress at this point can be either added up or canceled out. Since this stress directly affects the strain energy of a TSV structure, the ERR value also varies over different angles.

Now experiments with two aggressors located at  $10\mu m$  distance from the victim TSV are performed. Then, the angle among three TSVs are varied from  $45^\circ$  to  $180^\circ$  to explore the impact of angle on the ERR of the victim TSV as shown in Figure 67(a). As shown in Figure 68, the ERR is minimum when three TSVs form  $90^\circ$  angle and increases as the angle approaches to either  $45^\circ$  or  $180^\circ$ . If  $\sigma_{xx}$  stress component is plotted in Cartesian coordinate system, the TSV structure with a negative thermal load creates tensile stress along the x-axis and compressive stress along the y-axis. Thus, if a victim and aggressors form  $90^\circ$ , tensile and compressive stress from each aggressor TSV cancel out at the victim TSV location. This is why the lowest ERR is observed at  $90^\circ$ .



**Figure 67:** Simulation structure for angular dependency. Distance from victim to all aggressors is  $d$ . (a) Two aggressor TSVs. Aggressor 1 is fixed and aggressor 2 rotates. (b) Three aggressor TSVs. Aggressor 1 is fixed and aggressor 2 and 3 rotate.

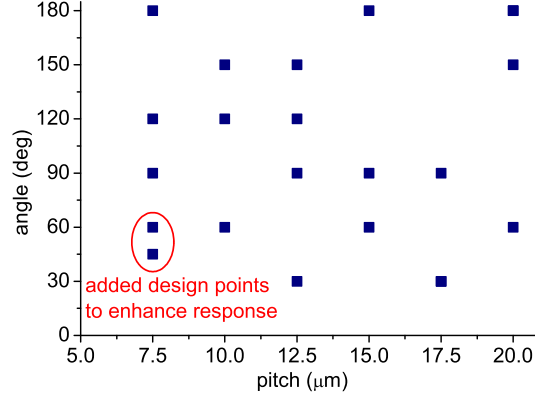


**Figure 68:** Impact of angle between victim and two aggressor TSVs on ERR.

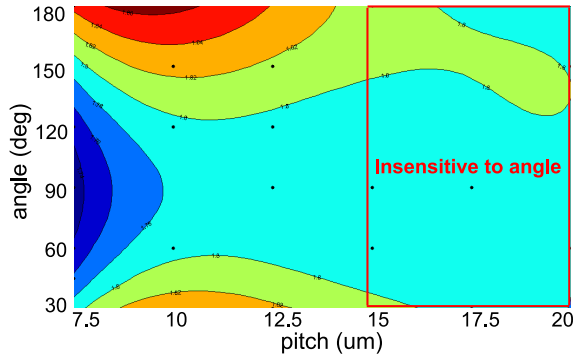
### 6.1.5 Relative Importance of Pitch over Angle

The combined effect of TSV pitch and angle is studied in this section. First, two aggressor TSVs are employed, and both pitch and angle are changed. In addition, DOE is utilized, which will be discussed in detail in Section 6.2, to produce simulation points as shown in Figure 69. Figure 70 shows an ERR contour map for different pitches and angles. High angular dependency is observed in a small pitch region. However, as the pitch exceeds  $15\mu\text{m}$ , the impact of angle is almost negligible.

To further investigate the relative importance between pitch and angle, three aggressor TSVs are used as shown in Figure 67(b). As shown in Figure 71, the angular dependency is dominant for small pitches. However, as the pitch increases, even though there are still some fluctuations along the angle axis, the angular dependency of ERR is not significant.



**Figure 69:** DOE based simulation points of two aggressor TSVs.



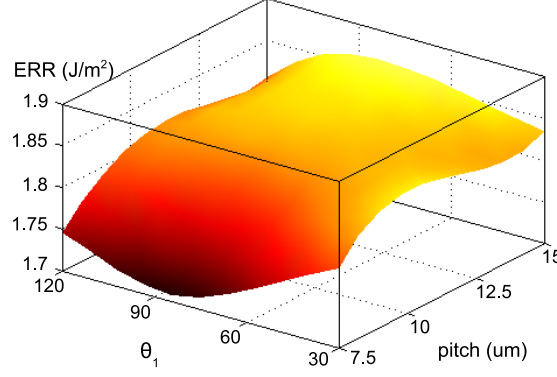
**Figure 70:** ERR contour map of pitch and angle.

Simulations with more aggressors up to eight are also performed. It is found that this angular dependency is almost not noticeable beyond  $10\mu m$  pitch, and the number of aggressor TSVs as well as the TSV pitch mostly determine the ERR value of the victim TSV.

## 6.2 DOE and RSM based Full-Chip TSV Interfacial Crack Modeling

FEA simulations for a TSV interfacial crack with multiple TSVs require huge computing resources and time. In the simulations, depending on the number of TSVs and mesh structure, a single FEA simulation takes from 1 to 12 hours using four CPUs. Thus, it is not feasible for full-chip-scale analysis.

Meanwhile, DOE has been used for many science and engineering applications. Recently, DOE was even used for co-optimization of power network, thermal TSV, and micro-fluidic channel in 3D ICs [59]. It has been proven to be an effective technique when analyses are desired for complex systems with multiple input factors. It provides a well-organized way



**Figure 71:** ERR surface map of pitch and  $\theta_1$  in three aggressor TSVs case.

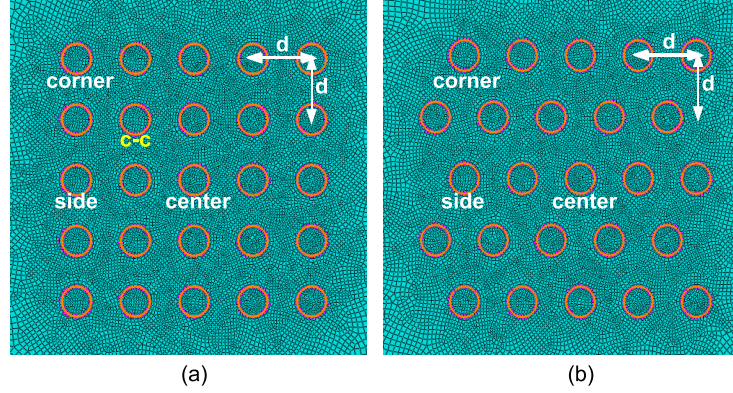
of performing experiments so that we can use the experimental results to find meaningful relations between input factors and responses of the system. In this section, a design of experiments (DOE) and response surface method (RSM) based full-chip TSV interfacial crack analysis flow is presented.

### 6.2.1 Designing Experiments

In general, TSV placement styles are largely divided into two categories: (1) Regular TSV placement. (2) Irregular TSV placement. To use DOE and RSM, design knobs (=input factors) and metrics (=responses) need to be defined. The ERR is used as a metric to assess TSV interfacial crack in a full-chip scale. However, input factors are different for regular and irregular TSV placement style.

Two possible regular TSV placement of  $5 \times 5$  TSV block with a same pitch is shown in Figure 72. In the array type, TSVs are aligned both in horizontal and vertical direction, whereas in the staggered type, TSVs in every other row are shifted by a half pitch. For both of these regular TSV placement schemes, the most important factor that determines the ERR of a victim TSV is the pitch and the position of the victim TSV inside the block such as center, side, or corner as shown in Figure 72. The TSV pitch is set as the only input factor and ERR models are built for aforementioned critical victim TSV locations separately for the regular TSV placement scheme.

Unlike the regular TSV placement style, there are countless possible combinations of TSV placement in the irregular TSV placement. However, simulating all these possibilities



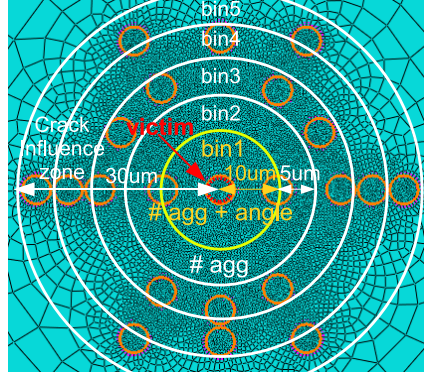
**Figure 72:** Top view of meshed simulation structures for  $5 \times 5$  TSV block. Orange circles are TSVs. (a) Array type. (b) Staggered type.

is impossible. The relative angle between a victim and an aggressor TSV is important only when the pitch is small as discussed in Section 6.1.5. In addition, as shown in Figure 65, the gradient of ERR along pitch is not steep. From these observations, the ERR model is simplified for the irregular TSV placement as follows: If the distance between a victim and aggressors is less than  $10\mu m$ , both the number of aggressors and angle between them are considered. If the distance exceeds  $10\mu m$ , only the number of aggressors at each distance bin with  $5\mu m$  interval is counted shown in Figure 73. In this way, the distance is removed from input factors, and the number of aggressors at each bin and the angle of nearest aggressors are used to design experiments.

Design points are generated using Stratified Latin Hypercube from space filling design styles. Based on the design points, FEA simulation structures are created and ERR data is obtained from each simulation. With these ERR values, the response surface and hence the compact ERR model are created for full-chip TSV interfacial crack analysis.

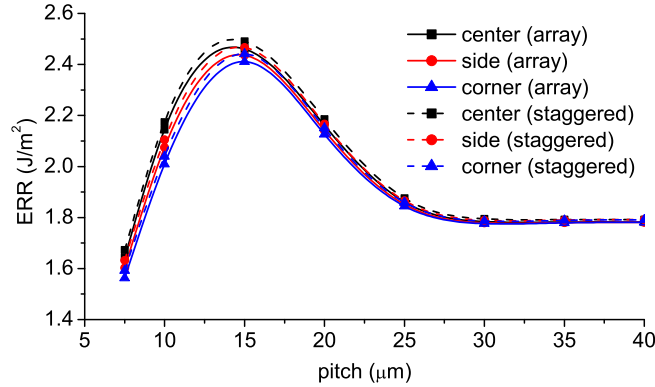
### 6.2.2 ERR Model for Regular TSV Placement

In this section, ERR models for the regular TSV placement are investigated. First, ERR values of victim TSVs in center, side, and corner locations shown in Figure 72 are monitored. As shown in Figure 74, the ERR is highest at the center, and decreases as the victim TSV location moves to side. The lowest ERR occurs at the corner due to the decreased number of aggressors surrounding the victim TSV. It is also observed that the ERR of victim TSVs



**Figure 73:** Top view of meshed simulation structure for irregular TSV placement with 21 aggressor TSVs.

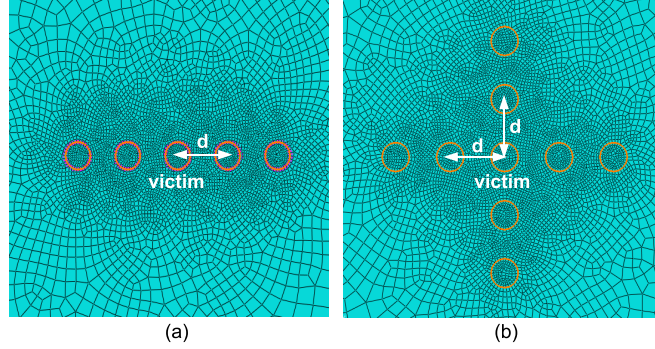
in the array type is always lower than the counterpart in the staggered type, even though the difference is not significant. This is because large numbers of nearby aggressors are forming  $90^\circ$  angle in the array type, hence reduces stress magnitude at the victim TSV location.



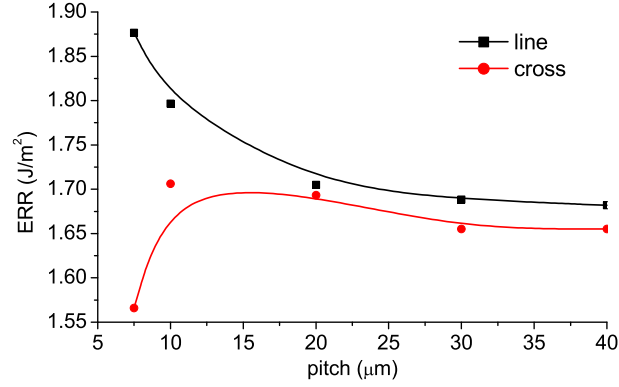
**Figure 74:** ERR vs. pitch for array and staggered type.

Interestingly and counterintuitively, the ERR is minimum at the smallest pitch and increases up to  $15\mu m$ , then it decreases and finally saturates at around  $30\mu m$  pitch. To verify why the minimum ERR occurs at the smallest pitch, two simulation structures are built as shown in Figure 75. In the line style, the victim TSV is only affected by the constructive stress interference from aggressors, whereas the victim TSV experiences both destructive and constructive stress interference from aggressors in the cross style. As shown in Figure 76, even though there are four more aggressors in the cross style, the ERR is always

higher in the line style for simulated pitches. Most importantly, the ERR is minimum at the smallest pitch in the cross style due to the highest destructive stress interference.



**Figure 75:** Top view of meshed simulation structures for (a) line style, (b) cross style.



**Figure 76:** ERR vs. pitch for line and cross type.

This observation indicates that it is always better to build TSV blocks in the array type rather than in the line style with a given number of TSVs to help suppress TSV interfacial crack growth. In addition, the ERR difference among three victim TSV locations, i.e., center, side, and corner, is higher in smaller pitches. This is again due to higher stress interference in smaller pitches, which results in larger stress magnitude differences among different victim TSV locations.

In this study, the array type is used for the regular TSV placement ERR model. Eight design points are generated for the regular TSV placement and an RSM model is built based on FEA simulation results. The RSM model can be expressed as a multivariate polynomial equation. In this case, the ERR model of the regular TSV placement is expressed as a 4<sup>th</sup>



order polynomial with one variable (=pitch) as follows:

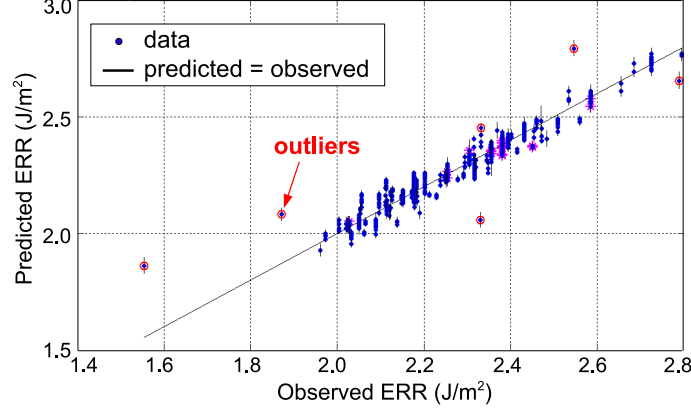
$$ERR_{reg} = c_1 + c_2 \cdot d + c_3 \cdot d^2 + c_4 \cdot d^3 + c_5 \cdot d^4$$

where,  $d$  is pitch and  $c_1 - c_5$  are TSV dimension dependent coefficients. ERR models for center, side, and corner locations are built separately. It is observed that the ERR of an intermediate point such as c-c shown in Figure 72(a) can be obtained by averaging ERR values of victim TSVs in center and corner locations with a negligible error. ERR models for different TSV array blocks are generated, such as  $3 \times 3$  and  $7 \times 7$  array.

### 6.2.3 ERR Model for Irregular TSV Placement

As discussed, there are innumerable scenarios for the irregular TSV placement. The number of input factors is reduced by distance binning and considering angular dependency within  $10\mu m$  distance from the victim TSV as shown in Figure 73. In addition, the number of aggressors at each distance bin cannot be arbitrarily large. Moreover, as the pitch becomes smaller, possible number of aggressors at that bin is also smaller due to the reduced bin area. 50 design points are generated with these constraints. Since 13 input factors are used, 8 for angle (8 aggressors are maximum possible number that can be placed in the bin 1) and 5 for number of aggressors at each distance bin, 50 design points are not enough to obtain a high quality RSM model. However, based on the observation that if the entire simulation structure is rotated by the same angle  $\theta$  around the victim TSV, the ERR will remain same since relative positions of TSVs are unchanged. Thus, 885 data points are generated from 50 simulations for better response surface fitting.

Figure 77 shows predicted ERR (RSM model) values and data points. There are 6 outliers which occur when there are no aggressors at bin 1, hence 8 input factors for angle are not exercised, which causes deviation from the predicted model. This ERR model can be enhanced by simulating more design points on this particular case. However, the proportion of the case with no aggressors at bin 1 is less than 6% (87 out of 1472 TSVs) in the worst case in our benchmark circuits. Moreover, since the ERR model fits well with data points in general, the model is used without further simulations. The ERR model of the irregular TSV placement is expressed as a  $2^{nd}$  order polynomial with 13 variables.



**Figure 77:** Predicted ERR using DOE and RSM vs. observed ERR.

#### 6.2.4 Quality of ERR Model

The goodness-of-fit of a model can be tested with statistics such as coefficient of determination ( $R^2$ ), root mean square error (RMSE), and prediction error sum of squares RMSE (PRESS RMSE) which is evaluated by excluding one data point at a time, building a new RSM model, and computing RMSE [59]. Table 33 shows that  $R^2$  values of the ERR models are close to 1, and both RMSE and PRESS RMSE is less than 0.1. Considering the fact that ERR values from the simulations range from 1.5 to 3.0 in general, the quality of fitting is acceptable.

**Table 33:** Quality of ERR model.

Placement type	$R^2$	RMSE	PRESS RMSE
Regular	0.993	0.034	0.086
Irregular	0.956	0.044	0.098

Even though the ERR models match well with simulation data, it is essential to validate whether our ERR models predict unseen data points correctly. Five new simulation structures are created to validate ERR models for both regular and irregular TSV placement cases. Table 34 shows predicted ERR values from our model and ERR values from simulations. Since the regular TSV placement type uses only one input factor, i.e., pitch, validation RMSE is lower than the irregular TSV placement case and closer to the model RMSE value. The validation RMSE of the irregular TSV placement type is also acceptable compared with the model RMSE value.

**Table 34:** Validation of ERR model. Simulation case shows pitch for regular TSV placement and number of aggressors for irregular TSV placement.

Placement type	Simulation case	ERR (model)	ERR (simulation)	Validation RMSE
Regular	9 $\mu m$	1.996	1.985	0.033
	12.5 $\mu m$	2.401	2.371	
	17.5 $\mu m$	2.355	2.335	
	22.5 $\mu m$	2.015	1.988	
	27.5 $\mu m$	1.789	1.778	
Irregular	10 agg	1.901	1.971	0.055
	21 agg	2.229	2.324	
	28 agg	2.320	2.305	
	36 agg	2.394	2.371	
	43 agg	2.572	2.547	

### 6.2.5 Full-chip Analysis Flow

Now, the full-chip TSV interfacial crack analysis flow is briefly explained. It is assumed that each TSV can be a candidate for a victim TSV. Thus, while each TSV is visited, this TSV is set as a victim TSV and other TSVs within influence zone as aggressors. In this simulation,  $30\mu m$  is used as a crack influence zone, since at around this pitch the ERR saturates as shown in Figure 74. Then, the angle and distance between an aggressor and the victim TSV are computed and the aggressor is inserted into the corresponding distance bin for the irregular TSV placement. For the regular TSV placement, the TSV pitch is found. Once this information is prepared, ERR values are computed using compact models based on DOE and RSM.

### 6.3 Full-Chip TSV Interfacial Crack Simulation Results

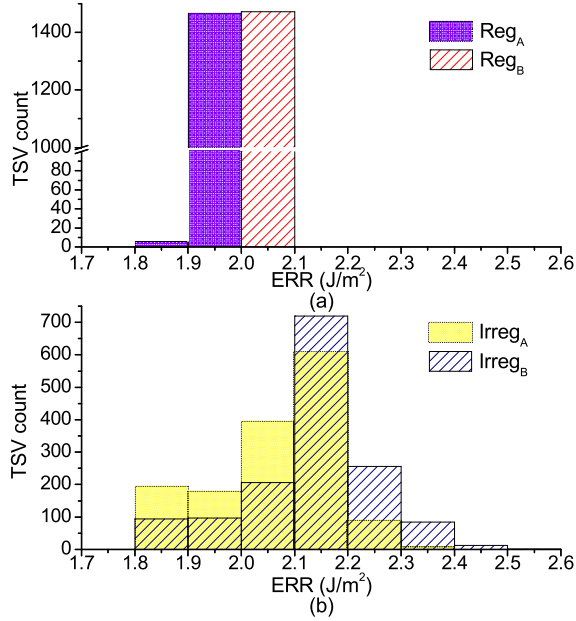
A full-chip TSV interfacial crack analysis flow is implemented in C++/STL. Four variations of a gate-level 3D circuit with changes in TSV placement style and TSV cell size are used for the analysis, which are listed in Table 35. The number of TSVs and gates are 1472 and 370K, respectively, for all cases. These circuits are synthesized using Synopsys Design Compiler [22] with the physical library of  $45nm$  technology [36] and designed to two-die stacked 3D ICs using Cadence Encounter [23].

**Table 35:** Benchmark circuits.

circuit	TSV placement	TSV cell size ( $\mu m \times \mu m$ )	wirelength ( $mm$ )	area ( $\mu m \times \mu m$ )	pitch ( $\mu m$ )
Irreg <sub>A</sub>	Irregular	$9.88 \times 9.88$	8884	$1000 \times 1000$	-
Reg <sub>A</sub>	Regular	$9.88 \times 9.88$	9648	$1000 \times 1000$	25
Irreg <sub>B</sub>	Irregular	$7.41 \times 7.41$	9060	$960 \times 960$	-
Reg <sub>B</sub>	Regular	$7.41 \times 7.41$	9547	$960 \times 960$	22

### 6.3.1 Impact of KOZ

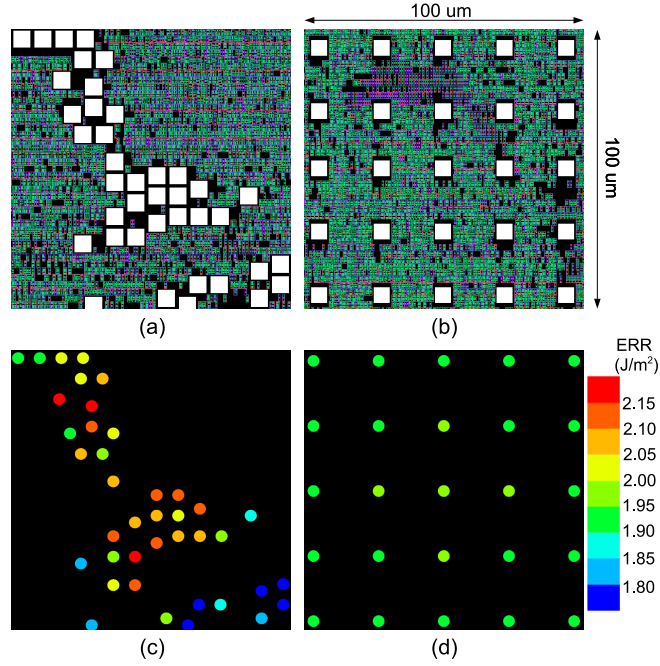
First, the impact of KOZ size on ERR of both regular and irregular TSV placement style is investigated. Figure 78 shows the histogram of number of TSVs for observed ERR ranges. It is found that ERR values are highly concentrated in a small range in the case of regular TSV placement. Even though there is a difference of ERR between center and corner locations in the TSV array for example, that is negligible in the pitch of  $22 \mu m$  (Reg<sub>B</sub>) and  $25 \mu m$  (Reg<sub>A</sub>). In addition, the KOZ size impact on ERR is not significant for this regular TSV placement case, since both TSV pitches are already close to crack influence zone ( $30 \mu m$ ), and their difference is only  $3 \mu m$ .



**Figure 78:** Impact of keep-out-zone on ERR. TSV<sub>A</sub> cell (KOZ =  $2.44 \mu m$ ) and TSV<sub>B</sub> cell (KOZ =  $1.205 \mu m$ ) (a) Regular TSV placement. (b) Irregular TSV placement.

### 6.3.2 Impact of TSV Placement Style

Next, the impact of TSV placement styles on ERR is investigated. In the irregular TSV placement case, a large variation of ERR values is observed. In addition, many TSVs experience higher ERR values than the regular TSV placement case. This is mainly because TSVs can be placed either densely or sparsely to minimize wirelength in the case of irregular TSV placement scheme. Thus, the ERR of a victim TSV can vary noticeably depending on the placement of nearby aggressor TSVs. Furthermore, since there are regions where group of TSVs are closely placed as shown in Figure 79(a), higher ERR values are observed in the irregular TSV placement style.



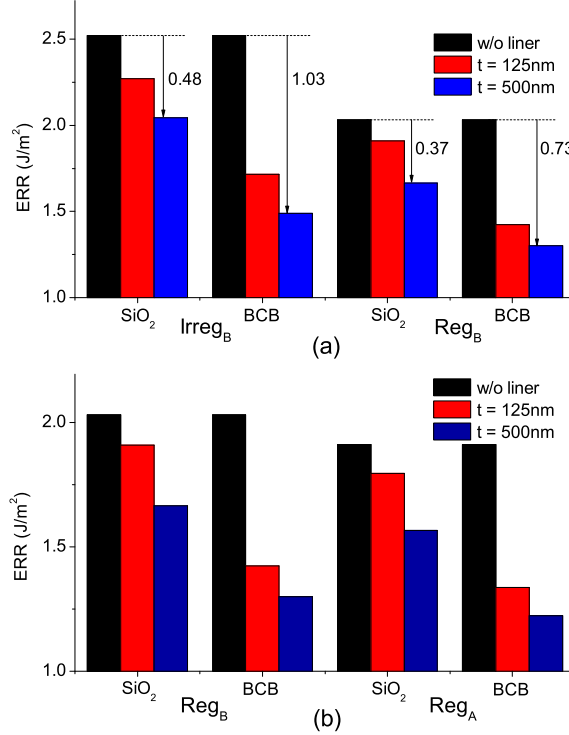
**Figure 79:** Close-up shots of layouts and ERR maps. (a) Irreg<sub>B</sub>. (b) Reg<sub>B</sub>. (c) ERR map of Irreg<sub>B</sub>. (d) ERR map of Reg<sub>B</sub>.

### 6.3.3 Impact of Liner

The ERR is highly dependent on the liner material and its thickness as discussed in Section 6.1.3. Therefore, the impact of liner on ERR in a full-chip scale is investigated. A  $6 \times 6 \mu m^2$  landing pad is used for all cases. Figure 80(a) shows that ERR values of both irregular and regular TSV placement schemes reduce significantly with a dielectric liner.

In addition, the liner thickness has a huge impact on the maximum ERR magnitude, since the thicker liner effectively absorbs thermo-mechanical stress at the TSV/liner interface. Especially, the BCB liner shows significant reduction in the maximum ERR compared with the SiO<sub>2</sub> liner due to extremely low Young's modulus. Furthermore, the ERR decrease is higher in the irregular TSV placement case compared with the regular TSV placement, since the stress buffer effect of liner is more effective where aggressor TSVs are close to the victim TSV.

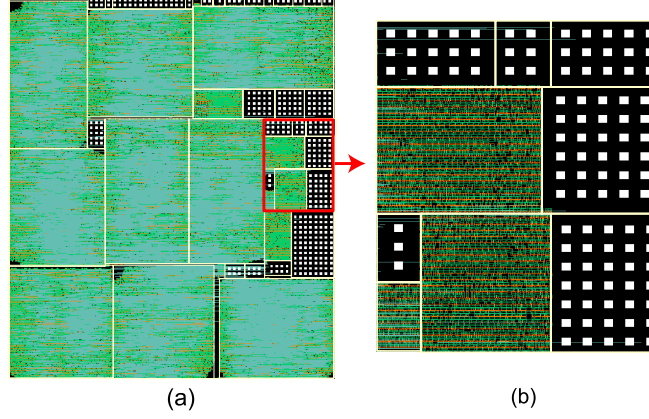
Figure 80(b) shows that differences of ERR values between Reg<sub>B</sub> and Reg<sub>A</sub> circuits are not significant with different liner material and its thickness. This is again because the pitch difference between Reg<sub>B</sub> and Reg<sub>A</sub> is negligible.



**Figure 80:** Impact of liner material and thickness on maximum ERR. (a) Irreg<sub>B</sub> vs. Reg<sub>B</sub>. (b) Reg<sub>B</sub> vs. Reg<sub>A</sub>.

#### 6.3.4 Reliability of Block-Level 3D Design

Even though the gate-level 3D design has the potential of highest optimization, the block-level design is attractive in the sense that highly optimized 2D IP blocks can be reused. In



**Figure 81:** Layout of block-level design (TSV pitch =  $15\ \mu m$ ). White rectangles are TSV landing pads. (a) full-chip layout. (b) close-up shot of the red box in (a).

this section, the TSV interfacial crack in block-level 3D designs is analyzed. 3D block-level designs are generated using an in-house 3D floorplanner which treats a group of TSVs as a block shown in Figure 81. A  $500nm$  thick BCB liner and a  $6 \times 6\ \mu m^2$  landing pad are used for all cases. The TSV pitch inside TSV blocks is varied to examine its impact on layout quality as well as reliability issues. Note that the pitch inside TSV block is smaller than regular TSV placement case, in general.

Table 36 shows that block-level designs use less number of TSVs, show shorter wire-length, and occupy more area than gate-level designs. Experimental results show that the block-level design with  $7.5\ \mu m$  pitch shows the smallest ERR among all cases. This is observed in Figure 74 due to strong TSV-to-TSV stress interference in small pitches. However, it comes with larger variation of ERR across TSVs compared with both irregular and regular TSV placement cases. This is mainly due to the small TSV pitch and different types of TSV blocks used in block-level design such as the line type as shown in top-right part of Figure 81(a). In addition, it is possible that the decreased TSV pitch could worsen signal integrity due to high TSV-to-TSV coupling. Therefore, the TSV pitch in block-level designs should be carefully determined considering both mechanical and electrical issues.

**Table 36:** Comparison between gate-level and block-level designs.

level	TSV pitch ( $\mu m$ )	# TSV	WL ( $mm$ )	area ( $\mu m \times \mu m$ )	max ERR ( $J/m^2$ )	std' dev'
Gate	irregular	1472	9060	$960 \times 960$	1.489	0.081
	22	1472	9547	$960 \times 960$	1.300	0.003
Block	7.5	333	7933	$980 \times 1090$	1.232	0.129
	10	394	8028	$1080 \times 1000$	1.500	0.160
	15	368	8259	$950 \times 1130$	1.805	0.270

## 6.4 Substrate Crack Modeling

### 6.4.1 Crack Propagation Modeling

The CTE mismatch-induced stress in 3D ICs may initiate cracks from the interface between a TSV and its dielectric liner, and propagates them on the silicon substrate surface. If a crack grows beyond the keep-out-zone (KOZ) of a TSV, it will jeopardize the reliability of the devices along its propagation path. In this section, the substrate crack is modeled to examine the crack propagation behavior under different scenarios such as an isolated TSV and TSV array.

#### 6.4.2 Isolated TSV Case

The simplest case is an isolated TSV. It is important to know how a crack will propagate in this case (in what direction and for how long). First, the direction of crack propagation is modeled, and it is shown that a crack is most likely to grow along the radial direction. Towards this, it is needed to prove that ERR attains maximum along the radial direction. It is known that the ERR is related to the stress intensity factor  $K_I$  as follows:

$$ERR = \frac{K_I^2}{E} \quad (19)$$

where  $E$  is the Young's modulus of material in which the crack propagates. Thus, it is needed to show that  $K_I$  attains maximum along the radial direction.

Consider the crack with given orientation as shown in Fig 82. The stress intensity factor  $K_I$  at the tip of the crack can be calculated as follows [33]:

$$K_I = r^2 B \Delta \alpha \Delta T \sqrt{\frac{\pi c}{8 R r^3}} \cos\left(\frac{\alpha}{2} + \frac{3(\alpha + \theta)}{2}\right) \quad (20)$$



where  $r$  is the TSV radius,  $B$  is the Young's modulus of the silicon substrate,  $\Delta\alpha$  is the CTE mismatch between TSV and silicon substrate,  $\Delta T$  is the thermal load,  $c$  is the crack length,  $\theta$  is the angle spanned by the crack from the center, and  $R$  is the outer radius of the  $n^+$  doped region. Applying trigonometry laws, following is obtained

$$\frac{r}{\sin \alpha} = \frac{c}{\sin \theta} = \frac{\sqrt{r^2 + R^2 - 2rR \cos \theta}}{\sin \theta} \quad (21)$$

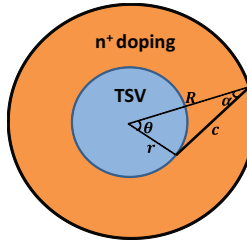
In addition,

$$\begin{aligned} & \cos\left(\frac{\alpha}{2} + \frac{3(\alpha + \theta)}{2}\right) \\ = & (1 - 2\sin^2 \alpha) \cos \frac{3\theta}{2} - 2\sin \alpha \cos \alpha \sin \frac{3\theta}{2} \end{aligned} \quad (22)$$

Substituting (21) and (22) into (20), we obtain

$$K_I = \frac{\cos \frac{\theta}{2} (4a\beta^2 \cos^2 \frac{\theta}{2} - 3a\beta^2 - 2\beta + 1)}{(\beta^2 + 1 - 2\beta \cos \theta)^{3/4} \sqrt{\beta}} \quad (23)$$

where  $a = \sqrt{r}B\Delta\alpha\Delta T\sqrt{\frac{\pi}{8}}$  and  $\beta = R/r$ . For  $\theta \in [0, \arccos(\frac{r}{R})]$  ( $\theta$  reaches maximum when the crack grows in the tangential direction of the TSV/liner interface circle) in (23), the denominator is increasing while the nominator is decreasing with the increase of  $\theta$ . Thus,  $K_I$  monotonically decreases with  $\theta$ , which leads to the following theorem.



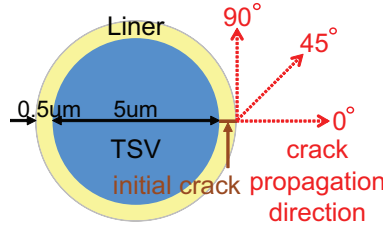
**Figure 82:** Top view of a crack ( $c$ ) initiated from TSV/liner interface and reaching the  $n^+$  doped region in arbitrary direction.

**Theorem 1:** For an isolated TSV, the crack along the radial direction has maximal  $K_I$  and ERR, while the crack along the tangential direction has minimal  $K_I$  and ERR.

Since the crack tends to grow in the direction with maximal ERR, the following corollary is obtained.

**Corollary 1:** For an isolated TSV, the crack grows and propagates along the TSV radial direction.

To validate Theorem 1, FEA simulation structures are built as shown in Figure 83. The TSV diameter, height, and dielectric liner thickness are  $5\mu m$ ,  $30\mu m$ , and  $0.5\mu m$ , respectively. Material properties used for the experiments are as follows: CTE (ppm/K) / Young's modulus (GPa) for Cu = (17/110), Si = (2.3/130), and SiO<sub>2</sub> = (0.5/71). The FEA simulation tool ABAQUS [37] is used to perform simulations, and all materials are assumed to be linear elastic and isotropic [60]. The entire structure undergoes  $\Delta T = +250^\circ C$  of thermal load (Cu electroplating  $25^\circ C \rightarrow$  annealing  $275^\circ C$ ) to represent a fabrication process.



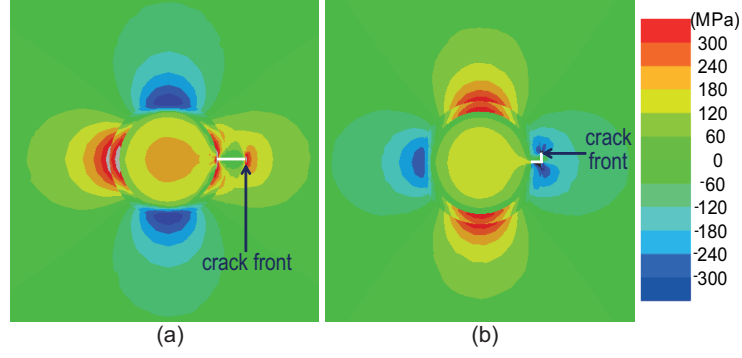
**Figure 83:** Top view of a crack simulation structure of a single TSV with different crack propagation directions. Initial crack length is  $0.5\mu m$  from TSV/liner interface.

The initial crack length and depth is assumed to be  $0.5\mu m$  and  $10nm$ , respectively, and this crack spans from the TSV/liner interface to the liner/substrate interface. Then, three cases are simulated to examine in which direction this crack will further grow as shown in Figure 83. ERR values from FEA simulations for these three cases are as follows: ERR ( $J/m^2$ ) for  $0^\circ$  (radial direction): 2.2,  $45^\circ$ : 0.6, and  $90^\circ$  (tangential direction): 0.4. This clearly shows that the crack will grow in the radial direction.

This preferred crack propagation direction, i.e., radial direction, can be explained with stress maps shown in Figure 84. For the crack to grow further, tensile stress needs to be applied perpendicular to the plane of the crack. With a positive thermal load ( $\Delta T = +250^\circ C$ ), Cu TSV expands more than silicon substrate, and hence generates compressive stress along the radial direction from the TSV center.<sup>1</sup> On the other hand, tensile stress builds up along the tangential direction around the TSV.

Thus, in the  $\sigma_{yy}$  map (Figure 84(a)), the tensile stress (red color) in the y-direction

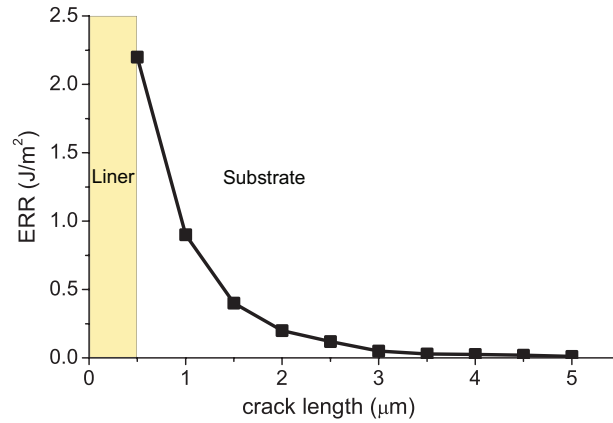
<sup>1</sup>With a negative thermal load (annealing  $275^\circ C \rightarrow$  room temperature  $25^\circ C$ ), tensile stress builds up around a TSV along the radial direction.



**Figure 84:** Stress maps around a crack front in an isolated TSV. (a)  $\sigma_{yy}$  map with crack propagating with  $0^\circ$  angle. (b)  $\sigma_{xx}$  map with crack propagating with  $90^\circ$  angle.

(tangential direction) opens the crack front that propagates in the radial direction ( $0^\circ$ ). On the other hand, when the crack propagates in the tangential direction ( $90^\circ$ ), the compressive stress (blue color) in x-direction (radial direction) closes the crack as shown in  $\sigma_{xx}$  map (Figure 84(b)).

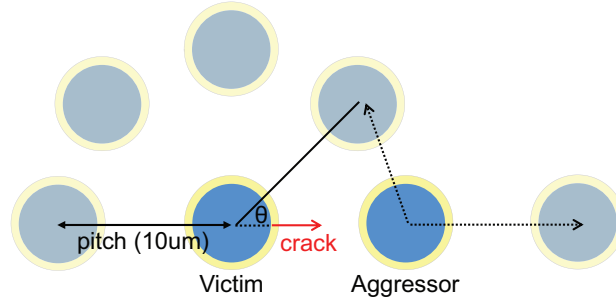
Now it is known that the crack grows in the radial direction, the next question is the maximum length it can grow. Figure 85 shows that ERR monotonically decreases as the crack grows. This is because stress magnitude decreases rapidly as the crack front moves away from the TSV, and hence the strain energy available for crack growth becomes smaller. Beyond  $4\mu m$  away from the TSV/liner edge, the ERR is almost zero, and hence the maximum length of the crack is around  $4\mu m$  for an isolated TSV.



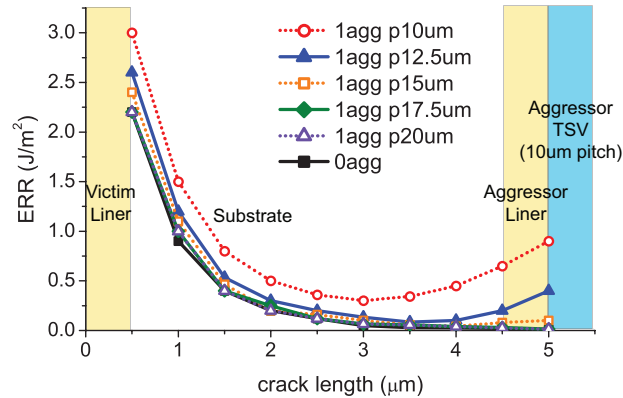
**Figure 85:** ERR of the crack propagating along the radial direction of a single TSV.

### 6.4.3 Single-Aggressor Case

In this section, the crack growth behavior is further investigated in the existence of a single aggressor TSV. Especially, the impacts of the TSV pitch (center-to-center distance) and location of the aggressor TSV with respect to the victim TSV are examined. The simulation structure is shown in Figure 86. It is assumed that the crack is growing along the x-direction (radial direction) and ERR values are obtained from FEA simulations. First, the aggressor TSV is placed with  $10\mu m$  pitch and then it moves away from the victim TSV up to  $20\mu m$  pitch with a  $2.5\mu m$  step. Figure 87 shows that as the crack propagates along the x-direction, the impact of an aggressor TSV with a pitch greater than  $15\mu m$  on the victim TSV crack's ERR is almost negligible (same as the case without an aggressor).



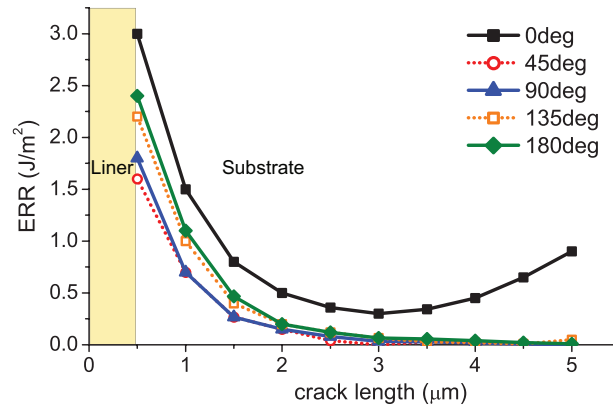
**Figure 86:** Simulation structure with one aggressor TSV. A crack grows along the x-direction (radial direction). The aggressor TSV either rotates around a victim TSV or moves away from the victim TSV.



**Figure 87:** ERR of the crack of a victim TSV with one aggressor at different pitches. Initial crack is  $0.5\mu m$  (from victim TSV/liner interface to liner/substrate interface). Crack length is swept from the initial crack up to the TSV/liner interface of the aggressor with  $10\mu m$  pitch.

However, interestingly when the TSV pitch is small enough such as below  $12.5\mu m$ , ERR values first drop as expected, and then start to increase as the crack front becomes close to the aggressor location. This is because the higher stress from the aggressor TSV creates higher strain energy for the crack to grow further. Thus, it is possible that a crack can create a bridge between two TSVs (bridge crack). As the copper atoms may easily migrate along the cracks, the bridge may create short-circuit issues.

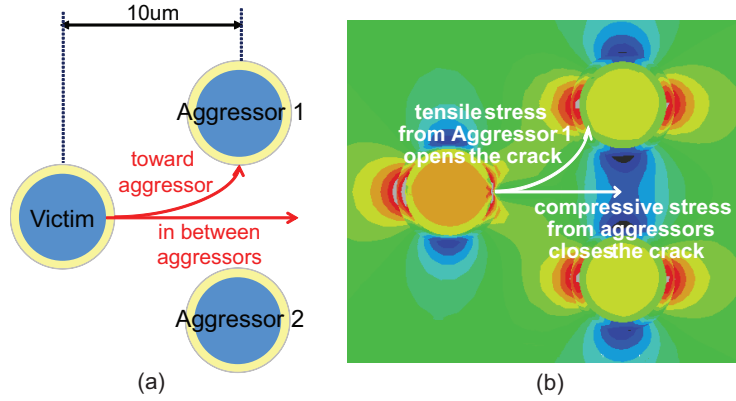
Now the aggressor at  $10\mu m$  pitch is rotated around the victim TSV with a  $45^\circ$  angle step. The ERR is highest when the aggressor TSV is on the crack propagation direction ( $0^\circ$ ) as shown in Figure 88. This is because the constructive stress interference between the victim and aggressor TSVs that generates higher tensile stress perpendicular to the crack propagation direction. Similar constructive stress occurs with an aggressor TSV at  $135^\circ$  and  $180^\circ$ , but that impact is much smaller than the  $0^\circ$  case because the distance between the crack front and the aggressor TSV is much longer and hence stress magnitude is lower. On the other hand, when the aggressor TSV is at  $45^\circ$  and  $90^\circ$  (acute angle), ERR values are lower than the other cases. In short, in the single aggressor case, the crack will either bridge the victim and aggressor TSV, or the crack will stop growing at certain distance similar to the single TSV case, depending the initial crack direction.



**Figure 88:** ERR of the crack of a victim TSV with one aggressor at different angles. TSV pitch is  $10\mu m$  for all cases.

#### 6.4.4 Two-Aggressor Case

In this section, the crack propagation direction is examined when two aggressor TSVs exist as shown in Figure 89. If a crack can propagate in between aggressor TSVs, it is possible that this crack break the entire chip in the worst case. However, as shown in Figure 89(b), the compressive stress from aggressor TSVs close the crack, and ERR of  $0J/m^2$  is observed in FEA simulations. Thus, the scenario that a crack propagates across a chip is unlikely to happen. On the other hand, when the crack propagates towards one aggressor TSV, the tensile stress along the tangential direction around the aggressor TSV helps open the crack further (non-zero ERR). Therefore, it can be inferred from these simulations that a crack will grow towards the aggressor TSV rather than propagating in between TSVs, and as such, the radial direction is no longer always the most likely one.

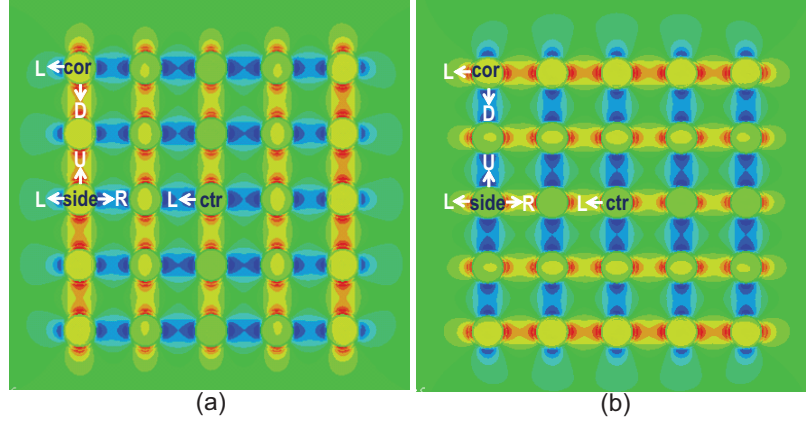


**Figure 89:** Crack propagation with two aggressor TSVs. (a) Simulation structure to examine if the crack propagates toward one aggressor or in between aggressors. (b)  $\sigma_{yy}$  map of (a).

#### 6.4.5 TSV-Array Case

Up to this point, the crack growth behavior is examined with up to two aggressor TSVs. In this section, the crack propagation in a general TSV array is investigated. ERR values of the crack of the victim TSV are monitored in three distinctive locations in a  $5 \times 5$  TSV array as shown in Figure 90: center, side, and corner locations. The pitch is still set as  $10\mu m$ . It can be easily inferred from stress maps that ERR values of cracks from any TSV other than these three locations will take intermediate values compared with these three

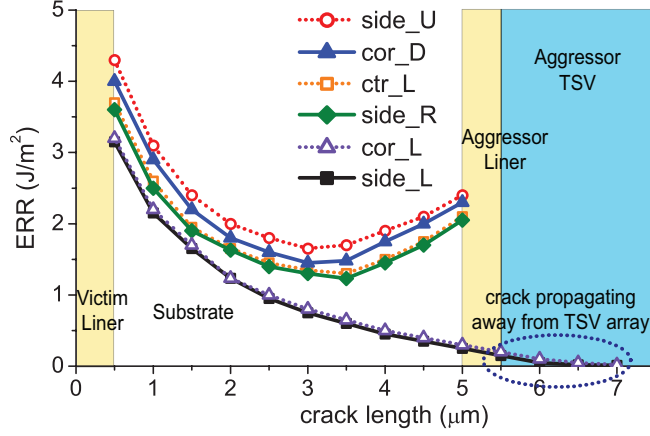
cases. Since it is observed that a crack growing toward an aggressor TSV has higher ERR in Section 6.4.2, the crack propagation direction is assumed to be upward, downward, left, or right from the victim TSV to show the worst case. All these directions are radial direction from the victim TSV.



**Figure 90:** Stress maps of a 5×5 TSV array. (a)  $\sigma_{xx}$  map. (b)  $\sigma_{yy}$  map.

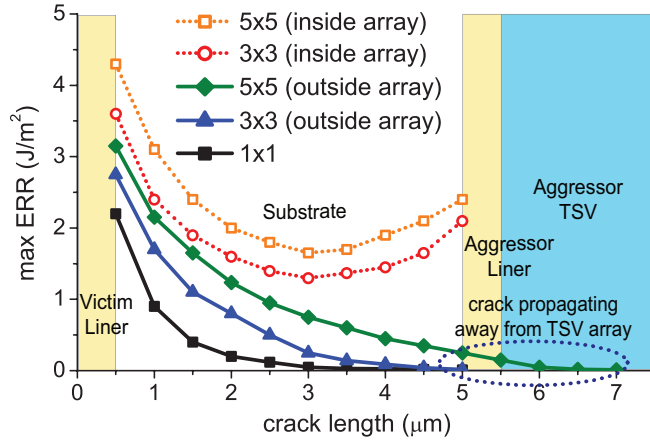
In addition, thanks to the symmetry in the array structure, only six cases need to be monitored. For example, as for the victim TSV in the center, ERR values for up, down, left, and right crack propagation directions are identical. Thus, only the crack that propagates to the left (ctr\_L) is monitored. As for the side location upward (side\_U), left (side\_L), and right (side\_R) directions are checked, and for the corner left (cor\_L) and downward (cor\_D) directions are tracked.

ERR curves for these six cases are shown in Figure 91. Interestingly, the ERR is highest in the side\_U case. As shown in Figure 90(a), the tensile stress in  $\sigma_{xx}$  map is highest around the side location along the upward and downward directions. Thus, this high tensile stress helps the crack to propagate in the upward direction. Although the victim TSV in the center is surrounded by more aggressor TSVs, the stress magnitude is lower than other locations because of the destructive stress interference between TSVs in vertical and horizontal directions. In all the cases, once a crack initiates from one TSV it is highly likely to grow and reach a neighboring TSV. On the other hand, for the crack growing away from the TSV array such as side\_L and cor\_L, the ERR decreases monotonically similar to the case without an aggressor, but decays much slower.



**Figure 91:** ERR of the crack of a victim TSV in different locations of a  $5 \times 5$  TSV array. side\_L and ctr\_L overlap. Naming convention: e.g., ctr\_L = a crack propagates towards left from the victim TSV in the center of a TSV array.

In addition, the impact of TSV array size on the crack propagation is studied:  $1 \times 1$ ,  $3 \times 3$ , and  $5 \times 5$  TSV arrays. The maximum ERR values of these three cases are monitored for the crack that grows inside TSV array (side\_U) as well as away from the array (cor\_L). As Figure 92 shows, larger TSV array generates higher ERR values. The conclusion here is that larger TSV array has higher probability of generating bridge cracks between two TSVs when they are growing inside the array, or cracks of longer length when they are growing away from the array.



**Figure 92:** Maximum ERR in different TSV array sizes. Inside array: a crack propagates inside a TSV array. Outside array: a crack propagates away from a TSV array.



## **6.5   *Conclusions***

In this chapter, impacts of TSV placement as well as TSV surrounding structures such as a liner and a landing pad on the TSV interfacial crack in 3D ICs is discussed. In addition, a DOE and RSM based accurate and fast full-chip TSV interfacial crack analysis flow is presented, which can be applicable to placement optimization for 3D ICs. Results show that KOZ size, liner material/thickness, and TSV placement are key design parameters to reduce the TSV interfacial crack problems in TSV based 3D ICs. In addition, detailed studies on the substrate crack propagation patterns under different scenarios are performed.

## CHAPTER VII

### LOW POWER DESIGN METHODOLOGIES FOR 3D ICS

Power reduction has been one of the most critical design considerations for IC designers. Minimizing both dynamic and leakage power is imperative to meet power budgets for portable devices (low power applications) as well as server farms (high power applications). The power efficiency also directly affects ICs packaging and cooling costs. In addition, the power of an IC has a significant impact on its reliability and manufacturing yield.

Because of the increasing challenges in achieving efficiency in power, performance, and cost beyond 32-22nm, industry began to look for alternative solutions. This has led to the active research, development, and deployment of thinned and stacked 3D ICs with TSVs. Black et al. studied the potential to achieve 15% power reduction as well as 15% performance gain of a high performance microprocessor by a 3D floorplan [16]. Kang et al. demonstrated 25% dynamic and 50% leakage power reduction in 3D DRAM [58].

In this chapter, physical design techniques that are shown to significantly reduce power consumption in 3D ICs are presented. This study is based on the OpenSPARC T2 design database [61] and a Synopsys 28nm PDK with nine metal layers that are both available to the academic community. GDSII-level 2D and 2-tier 3D layouts are built, analyzed, and optimized using the standard sign-off CAD tools.

A single OpenSPARC T2 core is first employed to demonstrate the 3D power benefit. First, how to rearrange functional unit blocks (FUB) into 3D to reduce power is discussed. Next, impacts of number of intra-block-level routing layers used on routing congestion and power consumption in 2D and 3D designs are examined. In addition, the impact of dual-Vth design technique on 2D and 3D power consumptions is presented. The effectiveness of functional unit block (FUB) folding, i.e., partitioning a FUB into two sub-FUBs and stacking them, in achieving power savings in the 3D design is also demonstrated.

Next, findings from the single core study are applied and extended to the entire OpenSPARC

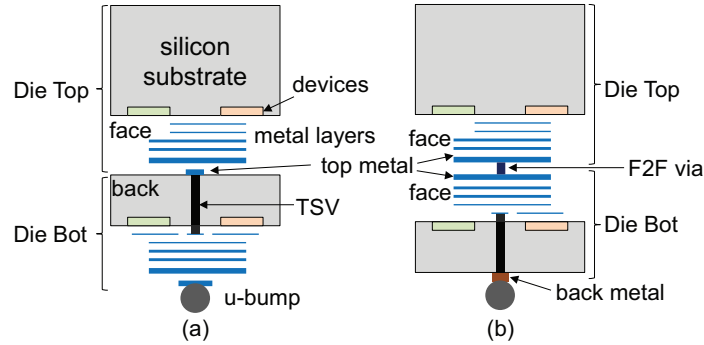
T2 system (an 8-core 64-bit SPARC SoC). In addition to aforementioned physical design methods, the impact of bonding styles, i.e., face-to-back (F2B) and face-to-face (F2F), on 3D power benefit is presented. For this, an efficient method to place face-to-face vias for the 2-tier 3D design utilizing existing commercial CAD tools with in-house scripts is developed.

## 7.1 Preliminaries

### 7.1.1 Die Stacking Technology

In this work, two-tier 3D ICs are employed. As shown in Figure 93, there are two possible bonding styles for 3D ICs: face-to-back (F2B) and face-to-face (F2F). In F2B bonding, TSVs are used for inter-die connections. Thus, the number of 3D connections can be limited by the TSV pitch as well as TSV area overhead. The face-to-face (F2F) bonding employing F2F vias is another attractive technology as this does not require additional silicon area for 3D connections.

The 3D interconnect settings are summarized in Table 37. TSV resistance and capacitance values are calculated based on the model in [27]. It is assumed that TSV diameter is much larger than F2F via size as manufacturing reliable sub-micron TSVs is challenging. Additionally, the physical size of F2F via can be made comparable to the top metal dimension, around twice the minimum top metal (M9) width in the Synopsys 28nm library.



**Figure 93:** Die bonding styles. (a) face-to-back. (b) face-to-face.

### 7.1.2 3D IC Design Flow

The RTL-to-GDSII tool chain for 3D IC design is based on commercial tools and enhanced with in-house tools to handle TSVs and 3D stacking. With initial design constraints, the

**Table 37:** 3D interconnect settings.

	diameter ( $\mu m$ )	height ( $\mu m$ )	pitch ( $\mu m$ )	R ( $\Omega$ )	C ( $fF$ )
TSV	3	18	6	0.043	8.4
F2F via	0.5	0.38	1	0.1	0.2

entire 3D netlist is synthesized. The layout of each die is done separately based on 3D floorplanning result. With a given target timing constraint, cells and memory macros are placed in each block. Note that only regular-Vth (RVT) cells are utilized as a baseline unless otherwise specified. The netlists and the extracted parasitic files are used for 3D static timing analysis (STA) using Synopsys PrimeTime [56] to obtain new timing constraints for each block’s I/O pins as well as die boundaries (= TSVs).

With these new timing constraints, block-level and chip-level timing optimizations (buffer insertion and gate sizing) as well as power optimizations (gate sizing) are performed using Cadence Encounter [23]. The design quality is improved through iterative optimization steps such as pre-CTS (clock tree synthesis), post-CTS, and post-route optimizations.

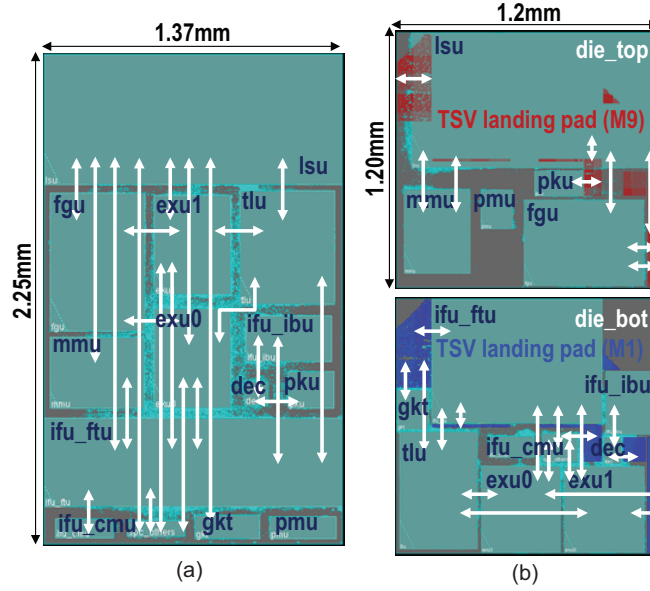
## 7.2 A Case Study with OpenSPARC T2 Core

### 7.2.1 3D Floorplanning Benefits

In this section, the implementation of both 2D and 3D block-level designs of OpenSPARC T2 core is discussed in detail. Then, based on layout simulations, several critical design metrics such as footprint area, wirelength, and power consumption of 3D designs with the traditional 2D designs under the same performance, i.e., iso-performance comparison, are compared.

The OpenSPARC T2 core consists of 13 FUBs including two integer execution units (EXU), a floating point and graphics unit (FGU), five instruction fetch units (IFU), and a load/store unit (LSU) [61]. Each FUB is synthesized with a 28nm cell library. In this implementation, top-level logic cells, i.e., cells outside FUBs, are grouped during synthesis to form an additional block. Thus, a total of 14 FUBs are floorplanned, and special cares are taken to use both connectivity and data flow between FUBs to minimize inter-block wirelength.

For the 3D design, the T2 core netlist is partitioned into two dies considering the area balance between dies and connectivity between FUBs. Then, the 3D floorplanner in [52] is employed with an objective of minimizing inter-block wirelength. In addition, two dies are assumed to be bonded in a face-to-back style. Note that TSV arrays are treated as additional blocks in this flow, hence all TSVs can be placed outside FUBs only. The total number of TSVs is 2979 in this design. The 2D and 3D placement results are shown in Figure 94. Note that intra-block (inside FUB) and inter-block routing utilize up to M5 and M9, respectively.



**Figure 94:** 2D and 3D placement results. (a) 2D design. (b) 3D design with 2979 TSVs. Cyan dots are core-level buffers. Blue and red rectangles are TSV landing pads at M1 and M9, respectively. White arrows represent major inter-block connections.

Comparisons between 2D and 3D block-level designs with a target clock period of 1.5ns (= 667MHz) are provided in Table 38. Note that these designs run much slower than UltraSPARC T2, a commercial product version of OpenSPARC T2, that runs at 1.4GHz [62]. This is mainly because some custom memory blocks in T2 core such as a content-addressable memory are synthesized with cells, since a general memory compiler cannot afford this kind of memories. Unfortunately, these synthesized memories are much larger and run slower than the memory macros generated by a memory compiler.

First, interestingly, the footprint area reduction in the 3D design is more than 50%.

**Table 38:** Comparison between 2D and 3D designs with a target clock period of 1.5ns. Numbers in parentheses are (intra block/inter block) breakdown.

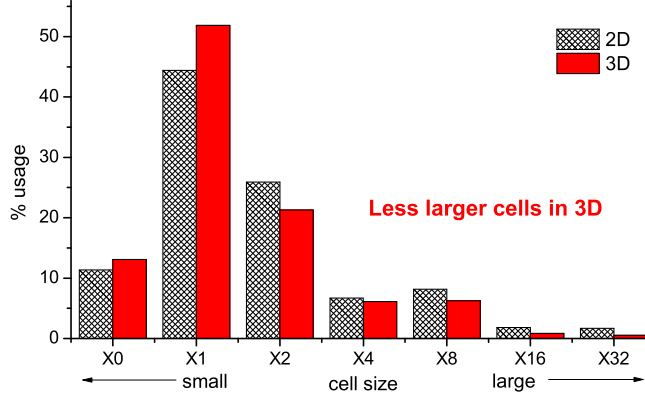
	2D	3D	diff
footprint ( $mm^2$ )	3.08	1.47	-52.3%
utilization (%)	67.8	66.8	-1.0%
# cells ( $\times 1000$ )	504.8 (483.8/21.0)	481.0 (458.8/22.2)	-4.7%
# buffers ( $\times 1000$ )	209.5 (188.5/21.0)	186.0 (163.8/22.2)	-11.2%
Wirelength (m)	23.3 (18.6/4.7)	20.2 (17.6/2.6)	-13.3%
<b>Total power (mW)</b>	<b>539.4 (489.3/50.1)</b>	<b>481.3 (455.5/25.8)</b>	<b>-10.8%</b>
Cell power (mW)	118.0 (111.9/6.1)	106.5 (100.8/5.7)	-9.7%
Net power (mW)	181.8 (150.9/30.9)	154.5 (137.3/17.2)	-15.0%
Leakage power (mW)	239.6 (226.9/12.7)	220.3 (213.3/7.0)	-8.1%

This is largely related to the buffer count reduction in the 3D design because of shorter wirelength and hence better timing. Note that the silicon area utilization, i.e., area occupied by cells, memory macros, and TSVs (3D only), for 2D and 3D designs are 67.8% and 66.8%, respectively, which supports a fair comparison.

Second, it is observed that 11.2% total buffer count reduction and 13.3% total wirelength decrease in the 3D design. However, counterintuitively, inter-block level buffers (= 22.2K) in the 3D design are more than the 2D (= 21.0K) even with the much shorter inter-block wirelength. As the design is optimized iteratively in FUB level and core level, buffers can be inserted either inside or outside FUBs to optimize paths. Additionally, to drive 3D nets with a large TSV capacitance, buffers need to be inserted. Thus, although inter-block level buffers are deployed more in the 3D design, a significant number of buffers is saved in the intra-block level. In addition, 5.4% intra-block wirelength reduction in the 3D design is observed mainly because of the intra-block level buffer count reduction.

Third, most importantly, the 3D design reduces power consumption over the 2D counterpart by 10.8%. It is found that cell (9.7%) and leakage (8.1%) power reduction are far more than the cell count decrease (4.7%) in the 3D design. As shown in Figure 95, the 3D design utilizes less larger cells than the 2D case thanks to better timing, i.e., more positive timing slack in paths. With the positive slack, cells can be downsized in the 3D design if this change still satisfies the timing constraint during power optimization stages.

This smaller cell size in the 3D design also helps reduce net power consumption. The



**Figure 95:** Cell size usage (%) comparison between 2D and 3D designs. X0 is the smallest cell size.

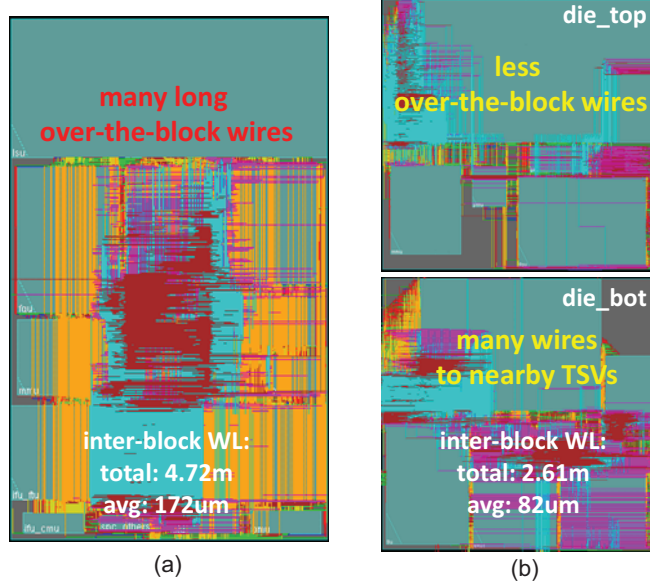
load capacitance of a driving cell is defined as the sum of wire capacitance and input pin capacitance of the loading side, hence the net power is defined as the sum of wire and pin power. Thus, the wire power reduction is directly from shorter wirelength, and the pin power decrease is from the smaller cell size as well as the reduced cell count.

### 7.2.2 Judicious Metal Layer Usage

So far, each FUB is routed using five metal layers to reserve sufficient routing resources for inter-block level routing that utilizes all nine metal layers. In this setting, four high metal layers can be used for over-the-block interconnections.

However, as shown in Figure 96, the inter-block routing demand is quite different between 2D and 3D designs. As for the 2D case, a large number of over-the-block wires are required, and this increases both total and average wirelength. Thus, more high metal layers (or global metal layers) are necessary to complete inter-block routing. On the other hand, many wires in the 3D design are connected to nearby TSVs, and this reduces over-the-block wiring demand significantly as well. Additionally, inter-block distance within a die is reduced with the reduced footprint area. As a result, the 3D design achieves a huge reduction in both total (44.7%) and average (52.3%) inter-block wirelength over the 2D design. Therefore, this 3D design may not need four high metal layers for over-the-block wiring.

Next it is investigated whether 3D design can provide further power saving by allowing



**Figure 96:** 2D and 3D inter-block routing results. Intra-block routing uses up to M5, and inter-block routing uses up to M9. (a) 2D design. (b) 3D design.

more metal layers for intra-block level routing. The key idea here is to reduce the amount of coupling capacitance inside FUBs by relaxing routing congestions with more metal layers and hence to reduce net power consumption. Three cases are studied in this work: intra-block routing up to M5 (baseline), M6, and M7.

The total power consumption of these three cases are shown in Table 39. All power numbers are normalized to the baseline 2D and 3D. As more metal layers are available for intra-block routing (less high metal layers for over-the-block wiring in inter-block routing), the 3D design further reduces power. For example, in the case of intra-block routing up to M7, the total wirelength and wire capacitance reduce by 1.4% and 3.5%, respectively, compared with the baseline. Note that the wire capacitance reduction is much more than the wirelength decrease, which indicates less routing congestion inside FUBs. This results in 5.8% net power and 2.9% total power saving.

**Table 39:** Impact of metal layer usage in intra-block level routing on power consumption for 2D and 3D designs. Power is normalized to the case of intra-block routing up to M5.

	intra-blk M5	intra-blk M6	intra-blk M7
2D power	1.0	1.026	1.021
3D power	1.0	0.977	0.971



However, in the 2D case, the opposite trend is observed largely because of the increase in both inter-block wirelength and buffer count. Moreover, the 2D design with intra-block routing up to M7 does not even close the target timing, and thus the power number is not reliable.

The impact of intra-block metal layer usage on intra-block and inter-block design metrics of the 3D design is shown in Table 40. We see that the 3D design with more intra-block metal layers achieves power reduction by improved intra-block level wirelength and buffer count that overwhelm the degraded inter-block level metrics.

**Table 40:** Impact of intra-block metal layer usage on intra-block and inter-block design metrics in the 3D design. Target clock period is 1.5ns and numbers in parentheses are difference with respect to the case of intra-block routing up to M5.

		intra-blk M5	intra-blk M6	intra-blk M7
Wirelength (m)	intra block	17.6	17.3 (-1.5%)	17.1 (-3.0%)
	inter block	2.6	2.7 (+2.7%)	2.8 (+8.8%)
	total	20.2	20.0 (-1.0%)	19.9 (-1.4%)
# buffers ( $\times 1000$ )	intra block	163.8	149.5 (-8.7%)	145.2 (-11.4%)
	inter block	22.2	22.9 (+3.2%)	25.9 (+16.7%)
	total	186.0	172.4 (-7.3%)	171.0 (-8.1%)
<b>Power (mW)</b>	intra block	455.5	443.5 (-2.6%)	439.7 (-3.5%)
	inter block	25.8	27.1 (+4.8%)	27.8 (+7.5%)
	<b>total</b>	<b>481.3</b>	<b>470.6 (-2.2%)</b>	<b>467.5 (-2.9%)</b>

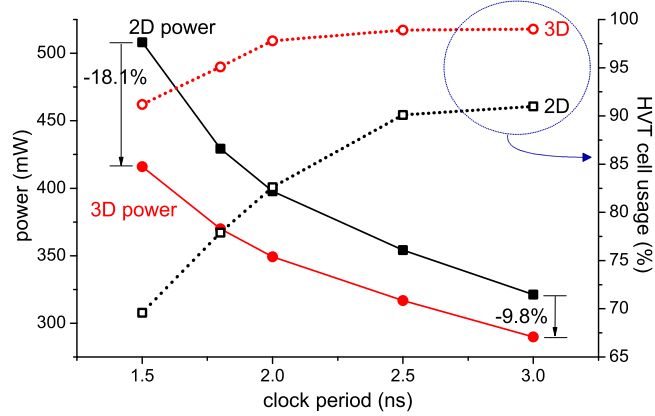
### 7.2.3 Dual-Vth Benefits for 3D ICs

Up to this point, both 2D and 3D designs utilize only regular-Vth (RVT) cells. However, industry has been using multi-Vth cells to further optimize power, especially for leakage power, while satisfying a target performance. In this section, high-Vth (HVT) cells are also employed to examine their impact on power consumption in 2D and 3D designs. Each HVT cell shows around 30% slower, yet 50% lower leakage and 5% smaller cell power consumption than the RVT counterpart.

To examine the 3D power benefit under different performances, five designs are implemented for both 2D and 3D cases: target clock periods are 1.5ns, 1.8ns, 2ns, 2.5ns, and 3ns. In all cases, a dual-Vth (DVT) cell library is utilized. As shown in Figure 97, 3D designs always use more HVT cells than 2D counterparts, and the HVT cell usage increases

as the target timing decreases. Even in the fastest case (1.5ns), the HVT cell usage in the 3D design is 91.2%, while that in the 2D design is only 69.6%. Thus, better timing in 3D designs translates to higher HVT cell usage, and this further reduces leakage power.

As shown in Figure 97, with a DVT design method, 3D designs benefit more in power reduction for faster cases. This is directly related to the HVT cell usage. At 1.5ns clock period, the 3D design reduces power consumption by 18.1%. As target clock period becomes slower, 2D designs also heavily utilize HVT cells and reduce the total power consumption noticeably, which decreases the 3D power benefit. Still, the DVT design method provides higher power improvement to 3D designs than RVT only cases for all target performances.



**Figure 97:** Power vs. delay curves and HVT cell usage for 2D (intra-block routing up to M5) and 3D (intra-block routing up to M7) DVT designs.

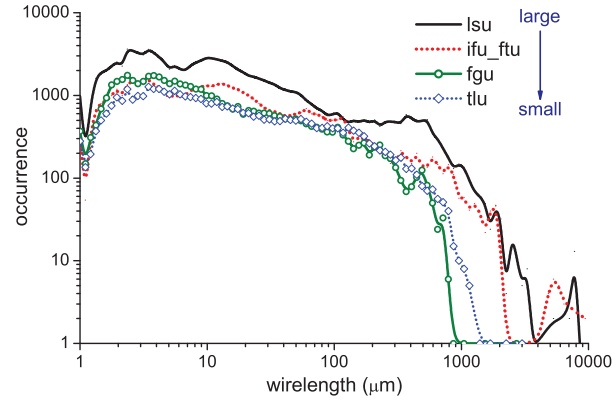
The DVT design technique reduces power noticeably for both 2D (5.8%) and 3D (13.6%) designs compared with the RVT only design at 1.5ns clock period. However, in the 2D case, the power saving is solely from leakage power reduction (17.6%). By employing weak HVT cells, the 2D design uses 7.6% more buffers and 6% longer wirelength than the RVT counterpart, which worsens cell and net power by 0.8% and 5.6%, respectively.

On the other hand, although the 3D DVT design uses slightly more buffers (1.5%) and longer wirelength (0.5%) than the 3D RVT design, cell power decreases by 4.2% since the HVT cell power is slightly lower than the RVT cell, and net power remains similar. Most importantly, leakage power decreases by 21.4%. Thus, the 3D design benefits more from the DVT design, especially for faster cases.

#### 7.2.4 Folding Functional Unit Blocks

So far, block-level designs are implemented for both 2D and 3D designs. Thus, even in 3D designs, each FUB is located in the same die. In addition, TSVs are always outside FUBs and used only for inter-block connections. In this section, the impact of FUB folding, i.e., partitioning a single FUB into two sub-FUBs and connect them with TSVs for intra-block connections, on power consumption is examined.

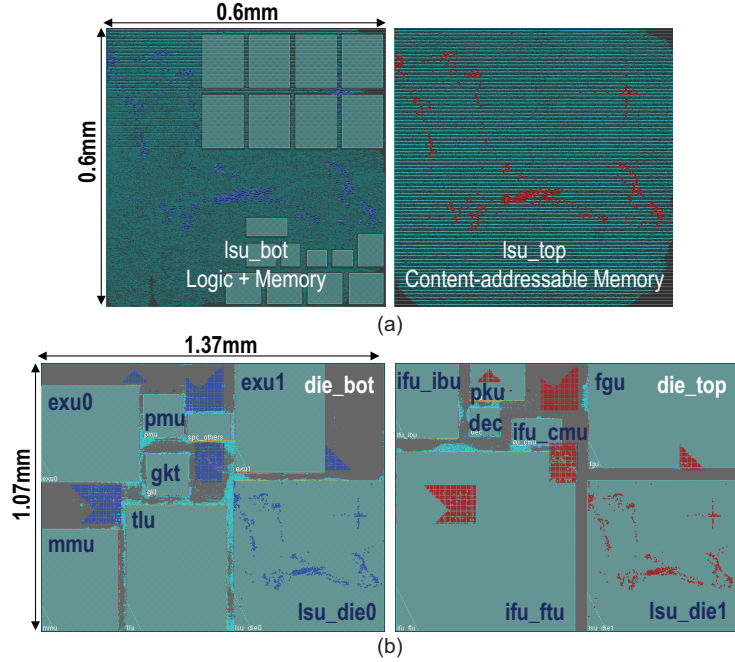
For the FUB folding to provide power saving, certain criteria need to be met. First, the target FUB needs to contain a large number of long wires so that wirelength decrease and hence net power reduction in the folded FUB can be nonnegligible. In general, large blocks tend to contain many long wires. Wirelength distributions of top four largest FUBs in the T2 core are shown in Figure 98. The top two largest FUBs, LSU and IFU\_FTU, are outstanding candidates.



**Figure 98:** Wirelength distributions of top four largest FUBs in T2 core.

Second, the target FUB is required to consume high enough portion of the total system power. Otherwise, the power saving from the FUB folding could be negligible in the system level. In this T2 core implementation, LSU and IFU\_FTU consume around 28% and 23% of the total T2 core power, respectively. Third, the net power portion of the target FUB needs to be high. If the FUB is cell and leakage power dominant, the wirelength reduction of the folded FUB may not reduce the total power noticeably. The net power portion of LSU, FGU, and TLU are about 33%, 47%, and 43%, while that of IFU\_FTU is only 17%. Therefore, in this T2 core case, LSU is the best choice for folding.

The LSU block is partitioned into two dies and designed with an in-house mixed-size 3D placer as shown in Figure 99. The dual-Vth design technique is also applied. This folded LSU block reduces the footprint, buffer count, and wirelength by 50.8%, 9.7%, and 7.1%, respectively, compared with the 2D LSU block. In addition, the HVT cell usage in the folded LSU is 96.8%, while that in the 2D LSU is 79.7%. More importantly, the total power of LSU is reduced by 5.4% largely due to the decreased net (9.2%) and leakage (4.9%) power.



**Figure 99:** Placement results of a folded FUB and a 3D block-level design with the folded FUB. (a) folded LSU block (TSV#: 596). (b) 3D design with the folded LSU (TSV#: 2411 (1815+596)).

Detailed comparisons between 2D, 3D without FUB folding (*3D w/o folding*), and 3D with FUB folding (*3D w/ folding*) designs are shown in Table 41. In *3D w/ folding*, the total power reduces by 21.2% compared with the 2D design and by 3.7% compared with *3D w/o folding*. Interestingly, the FUB folding helps both the folded block itself and the overall floorplan for power saving. The inter-block wirelength decreases significantly because of the increased flexibility of 3D floorplanning with smaller FUBs, i.e., the largest FUB is divided into two. In this design, inter-block wirelength decreases by 27.0%, which in turn reduces inter-block buffers by 29.6% compared with *3D w/o folding*. As a result, inter-block power

reduces by 23.6% compared with *3D w/o folding*.

**Table 41:** Comparison between 2D (intra-block routing up to M5), 3D without FUB folding (intra-block routing up to M7), and 3D with FUB folding designs with a target clock period of 1.5ns. Dual-Vth design technique is applied to all cases. Numbers in parentheses are difference against the 2D design.

	2D	3D w/o folding	3D w/ folding
footprint ( $mm^2$ )	3.08	1.47 (-52.3%)	1.47 (-52.3%)
utilization (%)	69.2	67.5 (-1.7%)	67.1 (-2.1%)
# cells ( $\times 1000$ )	532.3	471.9 (-11.3%)	450.9 (-15.3%)
# buffers ( $\times 1000$ )	225.5	173.6 (-23.0%)	157.5 (-30.2%)
# HVT cells ( $\times 1000$ )	370.4	430.4 (+16.2%)	444.8 (+20.1%)
Wirelength (m)	24.7	20.0 (-19.0%)	18.4 (-25.5%)
<b>Total power (mW)</b>	<b>508.2</b>	<b>416.0 (-18.1%)</b>	<b>400.7 (-21.2%)</b>
Cell power (mW)	118.9	100.6 (-15.4%)	99.1 (-16.7%)
Net power (mW)	191.9	145.0 (-24.4%)	135.6 (-29.3%)
Leakage power (mW)	197.4	170.4 (-13.7%)	166.0 (-15.9%)

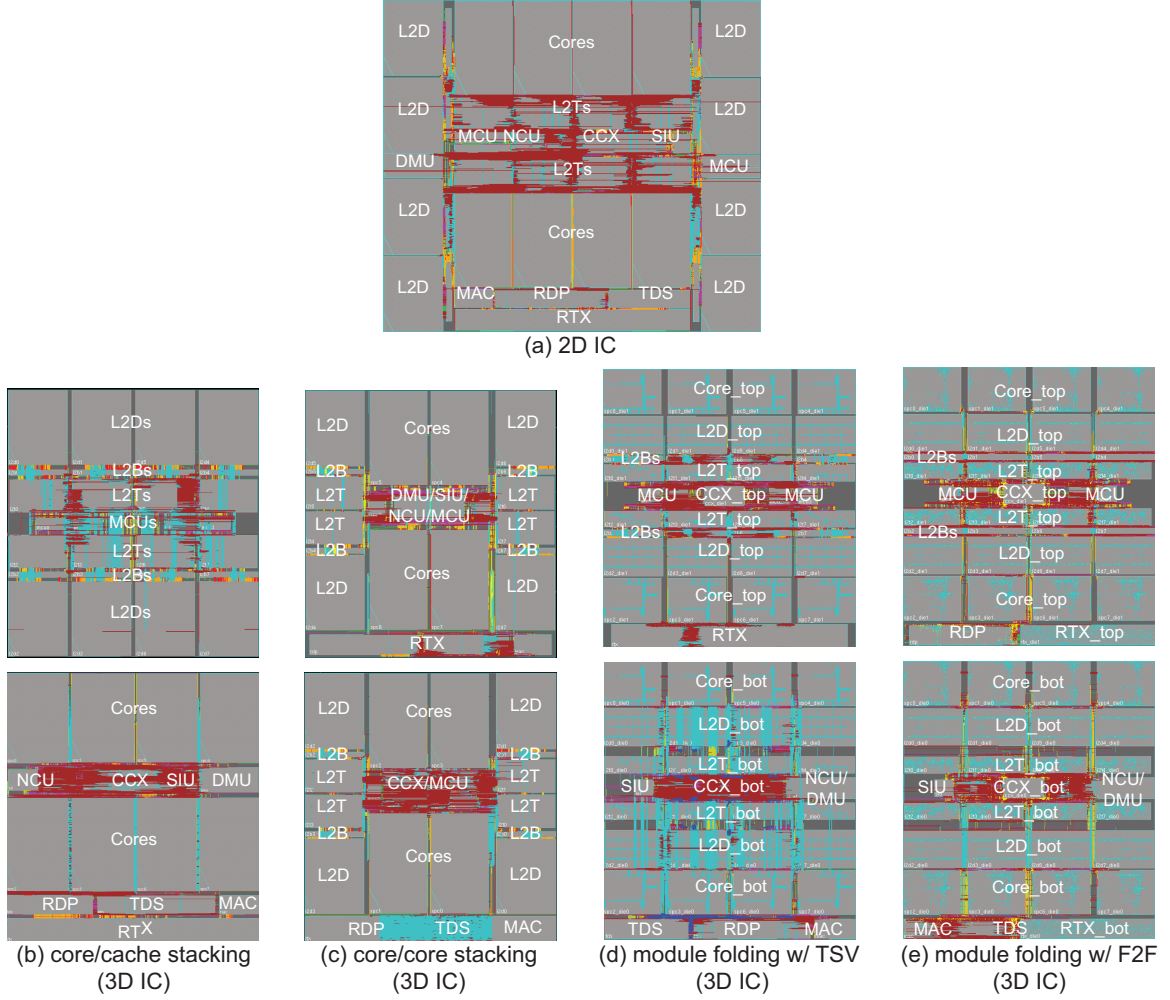
### 7.3 Low Power Computing Multi-core 3D Processors

In this section, all the lessons learned from the OpenSPARC T2 core case study are applied and extended in a much larger system. To further improve power benefits in 3D ICs on top of the aforementioned low power design methods, in-depth studies on impacts of block folding methodologies and die bonding styles are presented.

#### 7.3.1 Baseline Target System Design

The OpenSPARC T2, an open source commercial microprocessor from Sun Microsystems with 500 million transistors used, consists of 53 blocks including eight SPARC cores (SPC), eight L2-cache data banks (L2D), eight L2-cache tags (L2T), eight L2-cache miss buffers (L2B), and a cache crossbar (CCX). Each block is synthesized with 28nm cell and memory macro libraries. Seven blocks that do not directly affect the CPU performance are dropped from this implementation including five SerDes blocks, an electronic fuse, and a miscellaneous I/O unit. In addition, the PLL (analog block) in a clock control unit (CCU) is replaced by ideal clock sources. Thus, a total of 46 blocks are floorplanned. The 2D floorplan is designed to follow the original T2 floorplan [62] as much as possible as shown in Figure 100(a). In addition, special cares are taken to use both connectivity and data flow

between blocks to minimize inter-block wirelength.



**Figure 100:** GDSII layouts of 5 design styles of OpenSPARC T2 (full-chip) we compare: (a) 2D design ( $9 \times 7.9 \text{ mm}^2$ ), (b) core/cache stacking ( $6 \times 6.4 \text{ mm}^2$ ,  $\# \text{TSV} = 3,263$ ), (c) core/core stacking ( $6 \times 6.4 \text{ mm}^2$ ,  $\# \text{TSV} = 7,606$ ), (d) block folding with TSVs ( $6 \times 6.6 \text{ mm}^2$ ,  $\# \text{TSV} = 69,091$ ), (e) block folding with F2F ( $6 \times 6.6 \text{ mm}^2$ ,  $\# \text{F2F} = 112,308$ ). Cyan dots inside blocks are intra-block TSVs or F2F vias.

Note that all nine metal layers are utilized for SPC design that requires most routing resources among all blocks, but seven layers for all other blocks. Thus, top two metal layers can be utilized for over-the-block routing in the chip-level design.

The T2 chip contains eight copies of SPARC cores (SPC) and L2-cache blocks (L2D, L2T, and L2B) that occupy most of the chip area. These blocks need to be arranged in a specific order and a regular fashion for communication between them. Considering this constraint, area balance between dies, and connectivity between blocks, the T2 netlist is

partitioned into two dies. Two 3D floorplan cases to examine their impact on power as shown in Figure 100(b) and (c) are implemented: (1) core/cache stacking: all cores are in one die and all L2D blocks are in another die, and (2) core/core stacking: four cores and L2-cache blocks are located in each die.

The F2B bonding style is used for 3D block-level designs as a baseline. The 3D floorplanner in [52] is modified to handle user-defined floorplans, and then used to determine TSV locations with an objective of minimizing inter-block wirelength. TSV arrays are treated as additional blocks in this flow, hence all TSVs can be placed outside blocks only.

Now baseline 2D and 3D block-level designs are compared with a target CPU clock frequency of 500MHz that is the highest performance that our 2D design achieves. Design metrics in 2D and 3D designs are shown in Table 42. First, 16.3% buffer count and 5.0% wirelength reduction are observed in the core/cache 3D stacked design and 15.2% and 5.4% reduction in the core/core 3D case compared with the 2D counterpart. In addition, inter-block wirelength reduces by 15.6% (core/cache) and 17.8% (core/core), which is a direct consequence of 3D floorplanning.

**Table 42:** Comparison between 2D and 3D block-level designs with a target clock frequency of 500MHz. Numbers in parentheses are differences against the 2D design.

	2D	3D (core/cache)	3D (core/core)
footprint ( $mm^2$ )	71.1	38.4 (-46.0%)	38.4 (-46.0%)
# cells ( $\times 10^6$ )	7.39	7.21 (-2.4%)	7.26 (-1.8%)
# buffers ( $\times 10^6$ )	2.89	2.42 (-16.3%)	2.45 (-15.2%)
Wirelength (m)	343.0	326.0 (-5.0%)	324.5 (-5.4%)
<b>Total power (W)</b>	<b>9.107</b>	<b>8.171 (-10.3%)</b>	<b>8.273 (-9.1%)</b>
Cell power (W)	1.779	1.502 (-15.6%)	1.537 (-13.6%)
Net power (W)	4.499	4.122 (-8.4%)	4.131 (-8.2%)
Leakage power (W)	2.828	2.547 (-9.9%)	2.605 (-7.9%)

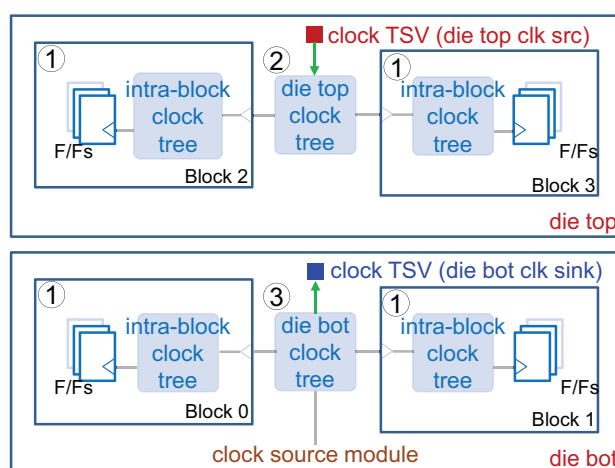
Second, most importantly, the 3D designs reduce power consumption over the 2D counterpart by 10.3% (core/cache) and 9.1% (core/core). The cell (15.6%) and leakage (9.9%) power reduction are far more than the cell count decrease (2.4%) in the core/cache 3D design. As discussed in the T2 core case, the 3D design utilizes smaller cells than the 2D thanks to better timing. This helps reduce cell, leakage, and net power consumption.

Third, the core/cache 3D stacking case shows 1.2% smaller power consumption than the

core/core case, which is essentially a negligible difference. This also indicates that there is not much room to further reduce power by 3D floorplans only, since there are not many floorplan options for the T2 design that contains multiple large same-size blocks that need to be placed in a specific way.

### 7.3.2 3D Clock Tree Router

One important tool that is currently missing in the EDA industry is 3D clock router. The 3D clock tree synthesis flow is briefly described in Figure 101. The goal is to use a commercial 2D clock router to build 3D clock trees while minimizing clock skew.



**Figure 101:** 3D clock tree synthesis flow (3D CTS): (1) Intra-block CTS, (2) Die top CTS. Clock TSVs are clock sources of die top. (3) Die bottom CTS. Die top clock tree information is utilized during die bottom CTS through clock TSVs.

First, an intra-block level clock tree is built for each block with given clock skew and slew constraints. Then, clock trees in die top are constructed, and clock TSV pads at M9 are clock roots in this case. In this step, the basic clock tree information of each block such as the minimum and maximum clock latency are transferred to the clock root pin of each block. This information is then utilized for skew control between blocks during the die top CTS. Finally, the die bottom CTS is performed where the clock source module is located. In this case, clock TSV pads at M1 are now clock sinks, and the clock tree information of die top is passed on these pads. Thus, during the die bottom CTS, the clock tree data of die top as well as blocks in die bottom are taken into account for both 2D and 3D clock

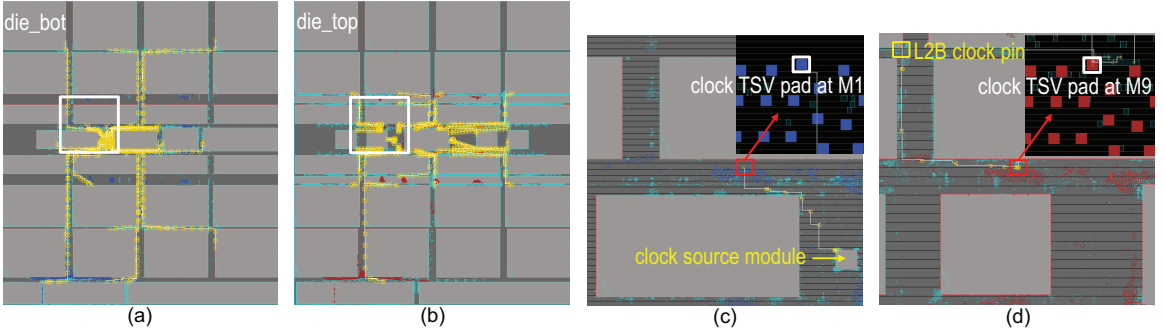


skew control.

2D and 3D CTS results are summarized in Table 43. The maximum clock skew is close to 12% of clock period for both cases. The 3D design uses 7.3% less clock buffer and hence 5.2% less clock power than the 2D design. The distance between the clock source block and clock pin of each block decreases in 3D, which in turn reduces clock buffer count. In addition, this clock source to each block's clock pin distance varies more in 2D than 3D due to the larger footprint area. This forces 2D design to use stronger clock buffers to balance clock skew across blocks. A snapshot of our 3D clock tree is shown in Figure 102.

**Table 43:** Clock tree comparison: 2D vs. 3D (core/cache).

	# clk buffers	max skew (ps)	clock power (mW)
2D	139.1K	243.6	768.5
3D	128.9K	233.1	728.5



**Figure 102:** Layouts of 3D clock tree of 3D design with four types of blocks folded (F2B). Yellow circles are clock buffers. (a) Die top clock tree. (b) Die bottom clock tree. (c) Close-up shot of white box in (a). (d) Close-up shot of white box in (b). 3D clock tree from clock source module (die bottom) to one L2B block (die top) is shown in (c) and (d).

#### 7.4 Block Folding Benefits

So far, block-level designs are implemented for both 2D and 3D designs. Thus, even in 3D designs, each block is located in the same die. In addition, TSVs are always outside blocks and used only for inter-block connections. In this section, the block folding method discussed in Section 7.2.4 is applied to blocks in the T2 system. In addition, the circuit characteristic specific block folding strategies are explored.

The block folding criteria are discussed in Section 7.2.4. First, the target block is

required to consume high enough portion of the total system power. Otherwise, the power saving from the block folding could be negligible in the system level. Blocks that consume more than 1% of the total system power are listed in Table 44. Note that the total power portion of SPC, L2D, and L2T is the average of corresponding eight blocks. Thus, SPC, L2D, and L2T are outstanding target blocks. In addition, RTX and CCX consume high power as a single block and hence could provide nonnegligible power benefit if folded.

**Table 44:** 2D design characteristics used for block folding candidate selection. Long wires are defined as wires longer than 100X standard cell height. CPU clock runs at 500MHz and I/O clock at 250MHz.

Block	Total power portion	Net power portion	# long wires	Remark
SPC	5.8%	55.1%	27.7K	CPU clock, 8X
RTX	3.6%	44.4%	27.5K	I/O clock
CCX	2.8%	57.6%	12.4K	CPU clock
L2D	2.1%	29.2%	6.5K	8X
L2T	1.8%	48.5%	6.0K	8X
RDP	1.7%	48.9%	5.2K	I/O clock
TDS	1.3%	43.1%	4.8K	I/O clock
DMU	1.1%	40.7%	5.4K	I/O clock

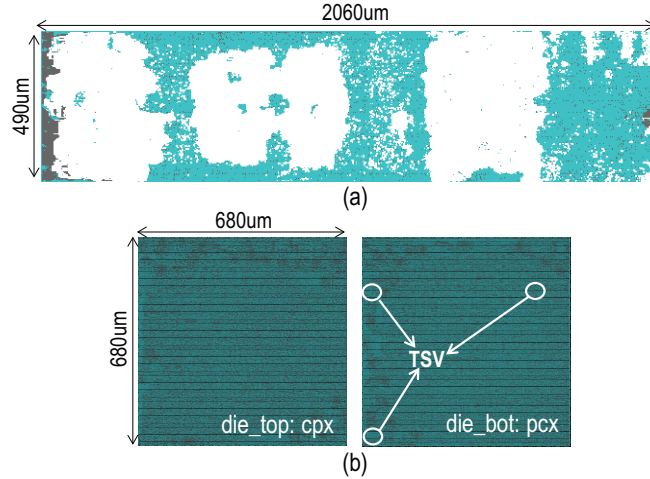
Second, the net power portion of the target block needs to be high. If the block is cell and leakage power dominant, the wirelength reduction of the folded block may not reduce the total power noticeably. Therefore, SPC and CCX are attractive blocks to fold. L2D shows relatively low net power portion compared with other blocks, as L2D is the memory (and its power) dominated design that contains 512KB (32 16KB memory macros in this implementation). Third, the target block needs to contain many long wires so that wirelength decrease and hence net power reduction in the folded block can be maximized. In this study, long wires are defined as wires longer than 100X standard cell height. The SPC, RTX, and CCX contain a large number of long wires.

In this work, five blocks are folded: SPC, CCX, L2D, L2T, and RTX. In the following sections, block folding methodologies for these five blocks are discussed. Each block shows distinctive folding characteristics.

### 7.4.1 Folding CCX Block

In T2, eight cores use the cache crossbar (CCX) to exchange data in eight L2-cache banks. This CCX is divided into two separate modules, the processor-to-cache crossbar (PCX) and the cache-to-processor crossbar (CPX). There are no signal connections between these two blocks except clock and a few test signals. The PCX occupies 48% of the block area and utilizes 48% of the CCX I/O pins, and the CPX uses the rest of them. Thus, the natural way to fold this CCX is placing the entire PCX block in one die and the CPX in another die along with related I/O pins.

The 2D and 3D CCX layouts are shown in Figure 103. Interestingly, in the 2D design, the PCX (and CPX) block is separated into several groups. The PCX has eight sources (SPCs) and nine targets (eight L2-cache banks and I/O bridge). Depending on the target core and L2-cache bank locations in the chip-level floorplan, PCX I/O pin locations are determined, which in turn attracts connected cells. Because of this, the PCX block is not fully gathered, which degrades cell-to-cell wirelength significantly.



**Figure 103:** CCX 2D and 3D layouts. (a) 2D design. CPX is highlighted with white color. (b) 3D design (# TSV=4).

However, folding CCX eliminates this problem and hence cell-to-cell wirelength decreases by 31.7% compared with the 2D as shown in Table 45. The folded CCX leads to 54.6% reduced footprint, 28.8% shorter wirelength, 62.5% less buffer count, and 32.8% power reduction over the 2D counterpart. Note that only four signal TSVs are used in

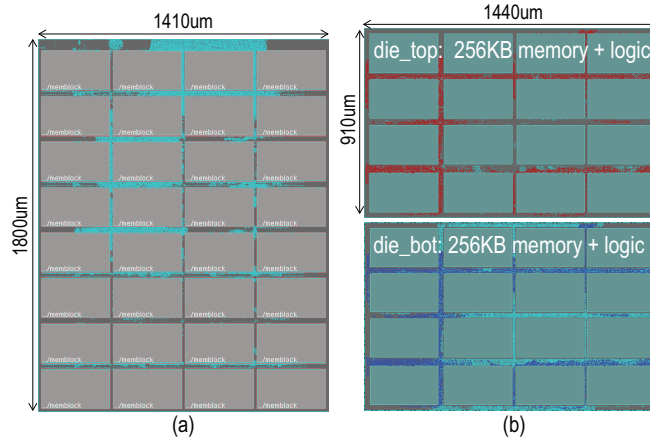
this 3D design, and this is due to the unique characteristics of CCX. It is also examined whether different 3D partitions with more 3D connections can provide better power savings. However, as the TSV count increases up to 6,393, largely due to the area overhead by TSVs (13.3%), the 3D power benefit reduces down to 23.4%.

**Table 45:** Wirelength comparison between 2D and 3D CCX.

Wirelength (m)	2D	3D	diff
I/O pin - cell	0.86	0.85	-1.2%
cell - cell	8.06	5.50	-31.7%
total	8.92	6.35	-28.8%

#### 7.4.2 Folding L2D Block

The single L2-cache data bank contains 512KB memory array. This L2D is further divided into four logical sub-banks. In this implementation, each sub-bank group is partitioned into eight blocks of size 16KB each. This L2D is a memory macro dominated design, and hence there are not many 3D partitioning options to balance area after folding. Thus, two sub-banks are placed in each die along with related logic cells as shown in Figure 104.



**Figure 104:** L2D 2D and 3D layouts. (a) 2D design. (b) 3D design (# TSV: 3,119). Blue and red rectangles are TSV landing pads at M1 and M9, respectively.

Although, the buffer count and wirelength reduce by 33.5% and 6.4%, respectively in the folded L2D, their impact on the total power saving is not significant (5.1% reduction over 2D) as shown in Table 46. This is because both cell and leakage power are dominated by memory macros, which 3D folding cannot help unless these memory macros themselves

are folded. Additionally the net power portion is only about 29% of the total power in 2D, and hence the small net power reduction in 3D does not lead to a noticeable total power reduction. Still, the footprint area reduction of 48.4% is nonnegligible and this might affect chip-level design quality.

**Table 46:** Comparison between 2D and 3D L2D designs.

L2D	2D	3D	diff
footprint ( $mm^2$ )	2.54	1.31	-48.4%
Wirelength (m)	3.41	3.19	-6.4%
# cells ( $\times 10^6$ )	53.1	42.2	-20.5%
# buffers ( $\times 10^6$ )	38.1	25.3	-33.5%
<b>Total power (mW)</b>	<b>172.9</b>	<b>164.0</b>	<b>-5.1%</b>
Cell power (mW)	25.8	24.6	-4.7%
Net power (mW)	50.5	44.5	-11.9%
Leakage power (mW)	96.6	94.9	-1.8%

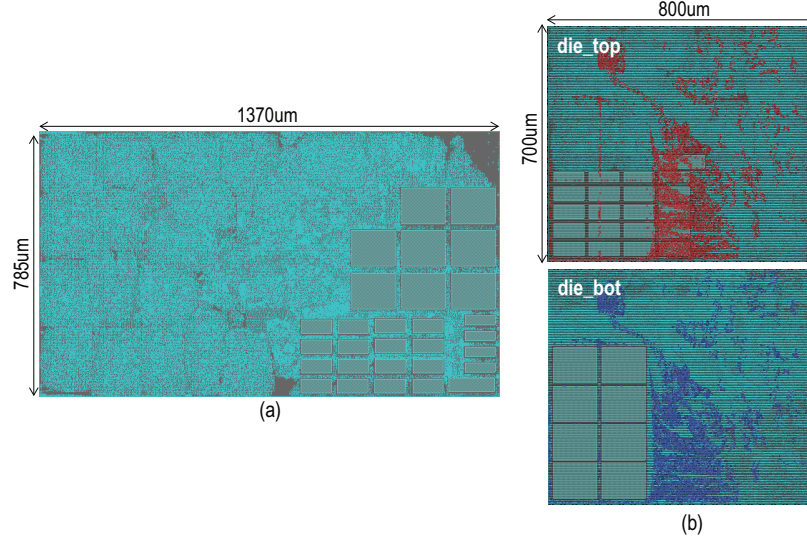
#### 7.4.3 Folding L2T Block

The L2-cache tag (L2T) consists of memory macros, synthesized memory blocks, and control logic cells in this implementation, and each of them occupies about one third of the total area. Partitioning options examined are listed in Table 47. Note that memory macros are divided into two groups based on their connectivities, i.e., tightly connected macros form a group.

**Table 47:** L2T die partitioning schemes.

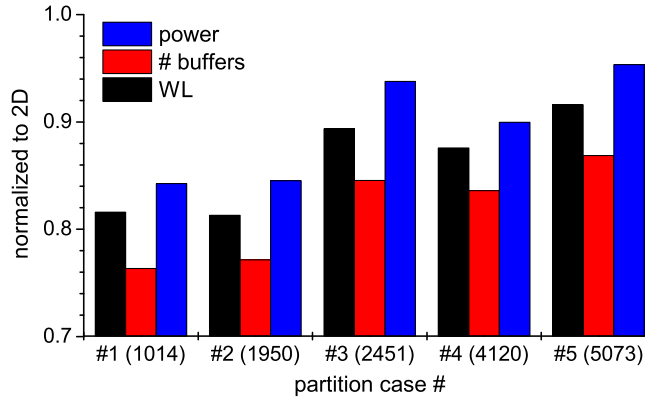
Part #	die bot	die top	# TSV
1	small macros syn' mem, logic	large macros syn' mem, logic	1014
2	small macros, logic	large macros, syn' mem	1950
3	syn' mem, logic	all macros, logic	2451
4	small macros, syn' mem	large macros, logic	4120
5	large macros, logic	small macros, syn' mem	5073

In partition #1, after splitting memory blocks, logic cells are partitioned using a min-cut partitioner, which leads to the smallest number of TSVs among five cases. On the other hand, in partition #5, where the largest number of TSVs are used, the silicon area occupied by TSVs is as high as 10% as shown in Figure 105. All these partitions are determined considering the area balance between dies including the TSV area.



**Figure 105:** L2T 2D and 3D layouts. (a) 2D design. (b) 3D design of partition #5 in Table 47 (# TSV: 5,073). The top die consists of small macros and synthesized memory. The bottom die contains large macros and logic cells. The total TSV area is 10%.

The die partitioning impact on the 3D design quality is shown in Figure 106. The partitioning cases with larger number of TSVs tend to lose the 3D power benefit. For example, partition #1 (# TSV: 1,014) shows 15.7% power saving, while partition #5 (# TSV: 5,073) achieves only 4.7% power reduction compared with the 2D. In these cases, the large TSV area overhead by TSVs results in increase in footprint area, wirelength, buffer usage, and hence power consumption. However, it cannot be generalized that more 3D connections degrade 3D design quality as this highly depends on 3D interconnect elements.



**Figure 106:** Impact of L2T die partitioning on wirelength, buffer count, and power. All numbers are normalized to the 2D.

#### 7.4.4 Folding RTX Block

The receiver-transmitter controller (RTX) block is the second largest block in this T2 implementation. The RTX consists of receiver controller (RXC) and transmitter controller (TXC), and shows similar characteristics as CCX. However, area is not balanced between two blocks: RXC (60% area) vs. TXC (40% area). Thus, it is not feasible to separate these two blocks into two dies.

Four different die partitioning options are implemented considering area balance between dies. The number of TSVs varies from 1,717 to 14,839 depending on the partition. A similar trend as in L2T is observed; a large number of TSVs degrades the 3D design quality. In the best case with 1,717 TSVs, RTX achieves 15.5% power saving over the 2D, while with 14,839 TSVs, only 7.9%.

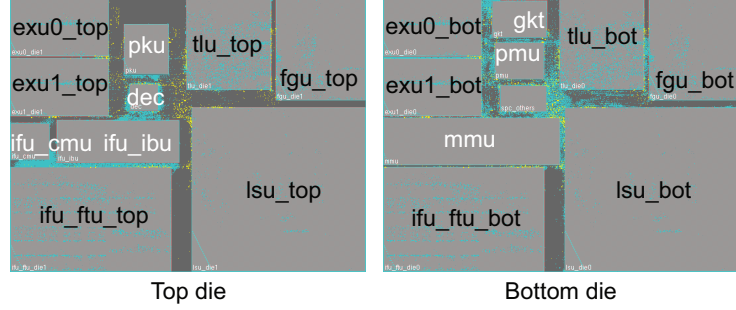
#### 7.4.5 Second-level Folding SPC Block

The SPARC core (SPC) is the highest power consuming block in T2. In Section 7.2.4, the 3D power benefit by folding the load/store unit (LSU) is demonstrated. In this section, more functional unit blocks (FUBs) are folded to further enhance the 3D power benefit. Since FUBs inside a SPC block are folded, we call this second-level folding.

Based on the block folding criteria, six FUBs are folded as shown in Figure 107. With this second-level folding, 9.2% shorter wirelength, 10.8% less buffers, and 5.1% reduced power consumption than the SPC without second-level folding, i.e. a block-level 3D design of the SPC, are achieved. Additionally, this 3D SPC achieves 21.2% power saving over the 2D SPC.

### 7.5 *Face-to-Face Bonding Benefit*

So far, 3D designs based on face-to-back (F2B) bonding using TSVs are discussed. In this section, it is examined that how face-to-face (F2F) bonding style utilizing F2F vias for 3D connections affects the 3D block folding design quality and power.

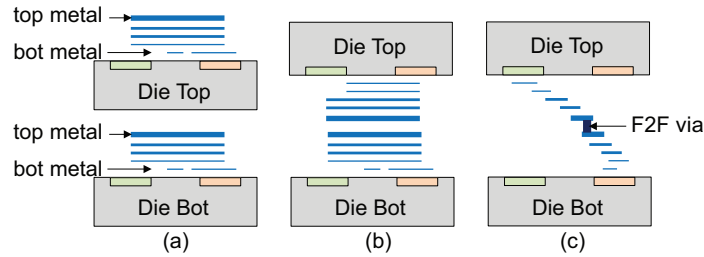


**Figure 107:** Second-level folding of a SPARC core. 6 FUBs shown in black text are folded (# F2F via: 10,251).

### 7.5.1 F2F Via Placer

Several previous works discussed TSV-aware 3D placement algorithms [63–65] assuming F2B bonding. However, there is no existing work on how to decide F2F via locations in F2F bonding. Unlike TSVs, F2F vias can be located above cells and macro blocks. Thus, 3D placement algorithms are not adequate for F2F via placement. In this section, the F2F via placement method by 3D net routing using existing commercial CAD tools is presented.

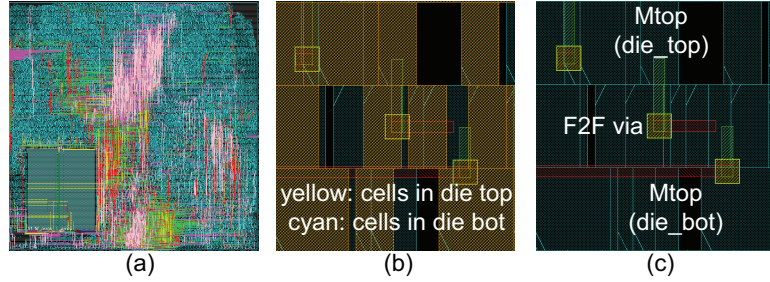
A simplified flow is shown in Figure 108. First, with a given die partitioning result, the 3D placer is run assuming an ideal 3D interconnect element (TSV size = 0) and then netlist and DEF (design exchange format) files for both dies are obtained. Next, 2D-like 3D design files, i.e., netlist, DEF, and LEF (library exchange format), are created that can be fed into commercial 2D place and route tools (in this case, Cadence Encounter). For example, 3D LEF file contains the interconnect structure for F2F bonding as well as cells and memory macros in both dies as shown in Figure 108(b). For this, metal layer and cell names are modified such as M1\_die\_top, M1\_die\_bot, INVX1\_die\_top, and INVX1\_die\_bot.



**Figure 108:** Finding F2F via locations by 3D net routing. (a) Run 3D placer assuming an ideal 3D interconnect. (b) Create 2D-like 3D design files (Verilog, LEF, and DEF). (c) Route 3D nets and extract F2F via locations.



Once all 3D design files are ready, a commercial CAD tool is employed to route 3D nets. In tool's perspective, these 3D nets are still 2D nets with cell pins located in either M1\_die\_top or M1\_die\_bot. Note that 2D net routing is excluded by modifying netlists: tying 2D nets to ground. By this, F2F via locations are not affected by 2D net routing and possible congestions. Layouts with 3D net routing and F2F via locations are shown in Figure 109. From this result, F2F via locations are extracted for all 3D nets, and these F2F via locations are utilized in each die design.



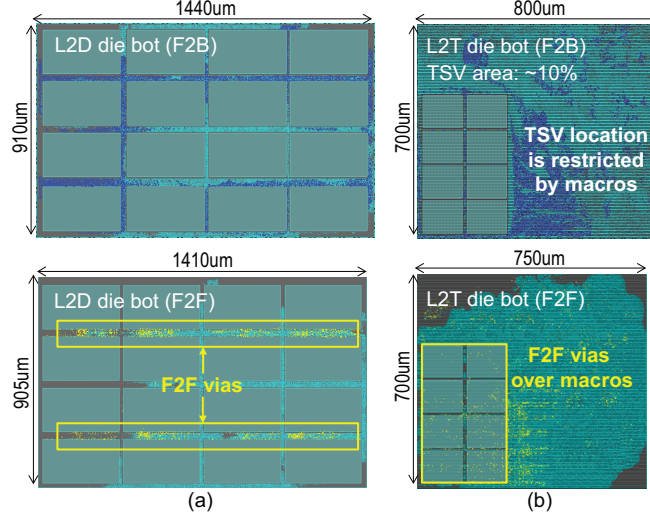
**Figure 109:** 3D net routing. (a) Layout shot after 3D net routing. (b) Close-up shot showing cells in both dies. (c) Close-up shot of 3D net routing showing F2F vias.

### 7.5.2 F2F Impact on Block Folding

F2F vias do not consume silicon area, and hence 3D footprint area can be further reduced as shown in Figure 110. For example, the folded L2D and L2T with F2F bonding reduce footprint by 2.6% and 6.3%, respectively, compared with F2B bonding cases. In the folded L2D case as shown in Figure 110(a), all F2F vias are located on horizontal channels between memory macros to connect memory I/O pins and logic cells right below them. On the other hand, TSVs are spread out all over the place because of their size and pitch. This affects cell placement as well, and hence degrades wirelength and power. For the same 3D partition, the folded L2D with F2F bonding shows 11.1% shorter wirelength, 3.9% less buffer count, and 4.1% less power consumption than the F2B case.

In addition, F2F via locations are not restricted by cells and macros. In the folded L2T case as shown in Figure 110(b), F2F vias are found over large memory macros. However, TSVs are ousted from memory macro area, which increases wirelength.

The five partitioning cases for L2T listed in Table 47 are implemented in both F2B

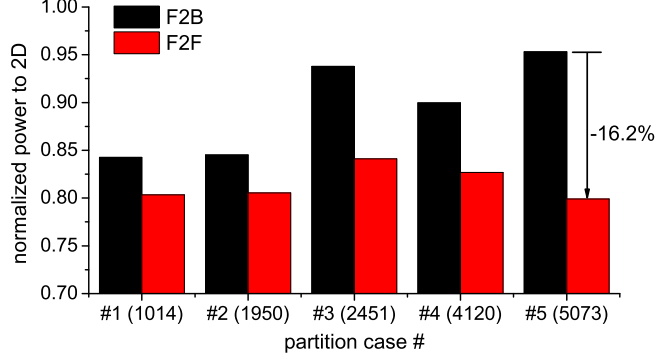


**Figure 110:** Bonding style impact on 3D placement. Blue rectangles are TSV landing pads at M1 and yellow dots are F2F vias. (a) L2D bottom die. (b) L2T bottom die.

and F2F bonding styles. Power comparisons between both bonding styles are shown in Figure 111. First of all, F2F wins over F2B bonding style in all cases. This is the combined effect of reduced footprint, better 3D connection points, shorter wirelength, less buffer usage, and better timing. Second, F2F bonding cases show larger power savings over the F2B cases in partition cases with more 3D connections. Especially, the partition #5 that shows the smallest 3D power benefit in F2B now achieves the best power saving with F2F bonding. Compared with the F2B case, the F2F case reduces power by 16.2%. In this specific case, the 3D design quality in F2B bonding is degraded largely by TSV area overhead, not by the partition. Third, more 3D connections in F2F style does not necessarily mean better power saving. Although partition #3 and #4 show much better power saving than the F2B cases, these power savings are still less than partition #1 and #2. This emphasizes the importance of die partitioning again.

## 7.6 Full-Chip Power Benefit With Folded Blocks

So far, impacts of block folding along with bonding styles on 3D power savings are discussed. In this section, all these folded blocks are integrated into a 3D T2 full chip and its impact on the system-level power is examined. In addition, the impact of bonding style on the full-chip 3D designs is presented.



**Figure 111:** Bonding style impact on power in L2T folding. Numbers in parentheses are number of TSVs/F2F vias.

### 7.6.1 3D Pin Partition

Another important CAD is needed for I/O pin partitioning for folded blocks, since inter-block routing quality is largely affected by block I/O pins. In the extreme case when all I/O pins of folded blocks are placed in the die bottom, routing congestion and detour will be serious in this die. This in turn increases coupling capacitance and thus net power consumption. This bad inter-block design quality can degrade intra-block design metrics as well. Therefore, I/O pins of folded blocks need to be partitioned so that inter-block wirelength in both dies is balanced.

As discussed in Section 7.4.1, CCX block is folded by placing PCX and related I/O pins in die bottom and CPX in die top. Input (output) pins of PCX are connected to SPCs (L2Ts), and the other way round for CPX. In this implementation, all SPC I/O pins are placed in the die bottom and L2T I/O pins in the die top to balance the number of inter-block wires related to CCX in each die. In addition, as L2-cache blocks are highly connected, all L2D I/O pins are placed in the die top. Note that only four TSVs are used in the folded CCX. One might expect this CCX folding incur a lot more inter-block TSVs between CCX and SPCs (or L2Ts). However, with aforementioned pin partitioning of SPC, L2T, and L2D, only 24 more inter-block TSVs are employed that are related to CCX compared with the 3D block-level design (= core/cache stack).

The same approach is applied to the RTX folding: MAC and TDS related I/O pins are placed in die bottom, and RDP related I/O pins in die top. With these pin partitioning

schemes, in five types of blocks folded case with F2B bonding discussed in Section 7.6.2, the inter-block wirelength is 11.7m and 10.1m for die bottom and die top, respectively. This is well balanced considering the fact that the distance between PCX (die bottom) and SPCs are longer than that between CPX (die top) and L2Ts as shown in Figure 100(d) and (e).

### 7.6.2 3D Floorplan with Folded Blocks

Based on the criteria on block folding discussed in Section 7.4, SPC, CCX, L2D, L2T, and RTX have been folded. Unlike other four blocks, RTX runs at I/O clock frequency (= 250MHz). In addition, almost all signals to/from RTX are connected with MAC, TDS, and RDP that form a network interface unit (NIU) with RTX. Thus, the impact of RTX folding is limited to the RTX block and NIU. In this study, two 3D designs are implemented as shown in Figure 100: (1) T2 with folded SPCs, CCX, L2Ds, and L2Ts, and (2) T2 with all five types of blocks folded.

In each case, two designs are built using either F2B or F2F bonding style. Note that there is a difference in routing layer usage in folded blocks depending on the bonding style. For the F2B bonding, the die bottom of folded blocks uses up to M7 (TSV landing pad at M1) as other unfolded blocks, while the die top utilizes up to M9 (TSV landing pad at M9). Thus, M8 and M9 can be used for over-the-block routing including folded blocks in the die bottom. The only exception is SPC that uses up to M9 for both dies as this block requires most routing resources. This is why SPCs are placed in top and bottom of the chip as shown in Figure 100(d) and (e). Otherwise, these SPC blocks will act as inter-block routing blockages.

In the F2F bonding case, since F2F via is on top of M9, all nine metal layers are used for routing. Thus, folded blocks in F2F bonding are inter-block routing blockages for both dies as shown in Figure 100(e). For this reason, although this F2F bonding achieves more power saving than the F2B case in block folding, inter-block design quality could be degraded.

In both bonding style cases, the CCX is placed in the center. There are about 300 wires between CCX and each SPC (or L2T). Thus, in this implementation, wires between CCX and L2T are much shorter than those between CCX and SPC. All other control units (SIU,

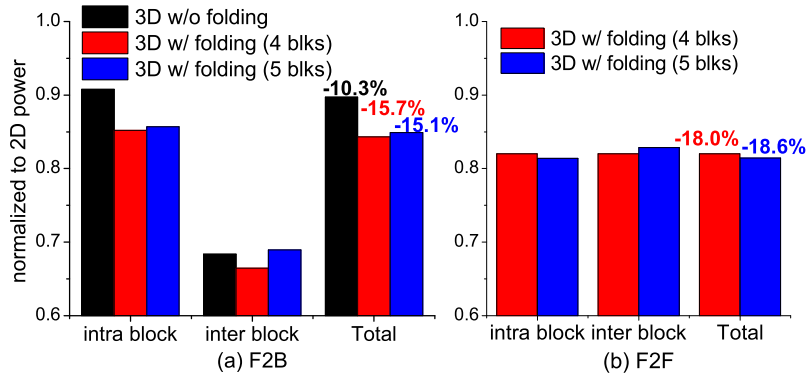
NCU, DMU, and MCU) are placed in the center row of each die as well. Finally, NIU blocks are placed in the bottom-most part of the chip as most of connections are confined in NIU.

### 7.6.3 Bonding Style Impact: F2B vs. F2F

As discussed in Section 7.5.2, block folding schemes (or die partitioning) are largely affected by bonding styles. For example, in L2T folding, the partition #1 (# TSV: 1,014) is the best for F2B, while partition #5 (# F2F via: 5,073) shows the lowest power for F2F. Thus, the best case for each block folding is selected depending on the bonding style, and then these folded blocks are integrated as shown in Figure 100(d) and (e).

In four types of blocks folded case, 69,091 TSVs (inter-block: 6,759, intra-block: 62,332) are used for F2B, and 101,555 F2F vias (inter-block: 6,759, intra-block: 94,796) are employed for F2F. In five types of blocks folded case, 69,123 TSVs (inter-block: 5,074, intra-block: 64,049) are used in F2B, and 112,308 F2F vias (inter-block: 5,074, intra-block: 107,234) are utilized in F2F. Note that by folding RTX, some of inter-block TSVs (F2F vias) are absorbed in intra-block TSVs (F2F vias).

3D T2 full-chip power normalized to 2D power is shown in Figure 112. First, more 3D power benefit is achieved with block folding (up to 18.6%) compared with the pure block-level 3D design (10.6%). It is also observed that the most of power saving is from intra-block level (= folded blocks). Note that the inter-block power is only about 5% of the total power.



**Figure 112:** Block folding impact on 3D full-chip power. (a) F2B bonding case. (b) F2F bonding case.

Second, inter-block power savings are worse in F2F cases. This is because all folded blocks act as inter-block routing blockages in the F2F bonding style, which increases inter-block wirelength and buffer count. For example, in four types of blocks folded case, inter-block wirelength and buffer count are 19.8m and 97.6K in F2B, respectively, while 22.5m and 122.3K in F2F. Third, however, power savings in intra-block level with F2F bonding overwhelm the loss in inter-block level, and hence F2F cases show better power benefit than F2B cases.

Lastly, as shown in Figure 112(a), folding more blocks does not always lead to more power saving. The RTX folding reduces inter-block power benefit in both bonding styles. Without RTX folding, connections between RTX and other NIU blocks are directly made by TSVs (or F2F vias). Thus, there are not many long horizontal wires as shown in Figure 100(d). However, with RTX folding, long horizontal wires are unavoidable between RTX and MAC, for example, as shown in Figure 100(e). In the F2B case, inter-block wirelength increases by 9.6% compared with the four block types folded case. This in turn degrades intra-block design quality as shown in Figure 112(a). Therefore, inter-block connections and its impact need to be considered when selecting blocks to fold.

#### 7.6.4 Full-chip Design Comparison with Dual-Vth Cells

Up to this point, both 2D and 3D designs utilize only regular-Vth (RVT) cells. As discussed in Section 7.2.3, 3D designs utilize more high-Vth (HVT) cells than 2D designs when dual-Vth (DVT) cell library is available. This is largely because of better timing in 3D.

First, three full-chip T2 designs are compared: 2D IC, 3D IC without folding (core/cache stacking, F2B bonding), and 3D IC with block folding (five types of blocks folded, F2F bonding), all with a dual-Vth (DVT) cell library. Detailed comparisons are shown in Table 48. As expected, 3D designs show higher HVT cell usage, especially for the 3D with folding case (94.0% of cells are HVT). This is largely due to better timing in 3D designs, and this helps reduce power in 3D ICs further. The 2D DVT design reduces power by 9.5% and the 3D with folding by 11.4% compared with the corresponding RVT only design, which again shows the benefit of 3D designs.

Most importantly, the 3D with folding case with F2F bonding reduces the total power by 20.3% compared with the 2D and by 10.0% compared with the 3D without folding case. This clearly demonstrates the powerfulness of block folding along with its bonding style in 3D designs for power reduction.

**Table 48:** Comparison between 2D, 3D without block folding (core/cache, F2B), and 3D with block folding (5 types of blocks folded, F2F) designs. Dual-Vth design technique is applied to all cases. Numbers in parentheses are difference against the 2D excluding HVT cell count which shows % of total cell count.

	2D	3D w/o folding	3D w/ folding
footprint ( $mm^2$ )	71.1	38.4 (-46.0%)	40.8 (-42.6%)
Wirelength (m)	339.7	321.3 (-5.5%)	309.6 (-8.9%)
# cells ( $\times 10^6$ )	7.41	7.09 (-4.3%)	6.83 (-7.8%)
# buffers ( $\times 10^6$ )	2.89	2.37 (-17.9%)	2.23 (-22.8%)
# HVT cells ( $\times 10^6$ )	6.50 (87.8%)	6.38M (90.0%)	6.42 (94.0%)
# TSV/F2F via	0	3,263	165,044
<b>Total power (W)</b>	<b>8.240</b>	<b>7.113 (-13.7%)</b>	<b>6.570 (-20.3%)</b>
Cell power (W)	1.770	1.394 (-21.2%)	1.175 (-33.6%)
Net power (W)	4.467	3.966 (-11.2%)	3.806 (-14.8%)
Leakage power (W)	2.003	1.753 (-12.4%)	1.589 (-24.2%)

Next, all five design styles of T2 full chip shown in Figure 100 are compared in terms of both intra-block and inter-block design metrics. Wirelength, buffer count, and power consumption are listed in Table 49. Note that, for 3D designs with block folding, the number of folded blocks is different for different bonding styles: Four types of blocks are folded for F2B, while five types for F2F. We choose the best case for each bonding style based on the results shown in Figure 112.

It is observed that the inter-block wirelength of 3D designs with block folding cases is worse than the 3D core/core stacking case. For example, the inter-block wirelength of 3D folding with F2F via case is 4.7% longer than the core/core stacking case. This is largely due to the fact that folded blocks are inter-block routing blockages, and hence inter-block routing detour is unavoidable, especially for the F2F bonding case. The same trend is observed in the inter-block buffer count and power.

However, block folding benefits are predominant in intra-block level metrics. The 3D folding with F2F via case achieves 8.4% shorter wirelength, 23.8% less buffer, and 19.9% smaller power than the 2D design in intra-block level. In addition, this F2F bonding case

shows 4.3% less intra-block power than the F2B case.

As discussed so far, the F2F bonding with F2F vias provides flexibility in die partitioning schemes for block folding as it does not require additional silicon area. This eventually leads to better system-level power saving in 3D if we choose proper target folding blocks, and then floorplan these blocks carefully considering connectivity with other blocks.

## **7.7 Summary**

In this chapter, the power benefit of 3D ICs was demonstrated with an OpenSPARC T2 chip. To further enhance the 3D power benefit on top of the conventional 3D floorplanning method, block folding methodologies and bonding style impact were explored. In addition, impacts of intra-block level metal layer usage control and dual-Vth design are examined.

An efficient method was also developed to find face-to-face via locations for 2-tier 3D ICs. It is demonstrated that more 3D power reduction is achieved with F2F bonding than F2B. With aforementioned methods, the total power saving of 20.3% has been obtained against the 2D counterpart in the OpenSPARC T2 full chip.



**Table 49:** Intra-block and inter-block level comparisons between 2D and 3D designs shown in Figure 100. Dual-Vth design technique is applied to all cases.  $\Delta$  is the difference against the 2D design.

		2D	3D (core/cache)	$\Delta$	3D (core/core)	$\Delta$	3D folding w/ TSV	$\Delta$	3D folding w/ F2F	$\Delta$
Wirelength (m)	intra block	316.2	301.0	-4.8%	304.4	-3.7%	294.5	-6.9%	289.7	-8.4%
	inter block	23.5	20.3	-13.6%	19.0	-19.1%	19.4	-17.4%	19.9	-15.3%
	total	339.7	321.3	-5.4%	323.4	-4.8%	313.9	-7.6%	309.6	-8.9%
# buffers ( $\times 10^6$ )	intra block	2.780	2.260	-18.7%	2.381	-14.4%	2.188	-21.3%	2.119	-23.8%
	inter block	0.109	0.111	+1.8%	0.085	-22.0%	0.105	-3.7%	0.112	+2.8%
	total	2.889	2.371	-17.9%	2.466	-14.6%	2.293	-20.6%	2.231	-22.8%
power (W)	intra block	7.829	6.831	-12.7%	7.159	-8.6%	6.556	-16.3%	6.273	-19.9%
	inter block	0.411	0.282	-31.4%	0.221	-46.2%	0.257	-37.5%	0.297	-27.7%
	total	8.240	7.113	-13.7%	7.380	-10.4%	6.813	-17.3%	6.570	-20.3%

## CHAPTER VIII

### CONCLUSIONS

As demonstrated in this dissertation and other works, 3D ICs provide significant benefits over traditional 2D ICs in important metrics such as footprint, wirelength, timing, power, and so on. Currently, industry is taking slow steps towards 3D IC because of various issues such as manufacturing cost, yield, thermal issues, logistics, lack of standards, etc. However, with the physical limits in devices and interconnects approaching fast, industry will eventually move towards 3D IC technologies. To successfully adopt 3D IC technologies, it is essential (1) to study the benefits of 3D IC designs based on today's and future technology settings, as well as (2) to develop the design methodologies for 3D ICs that resolve reliability problems (thermo-mechanical stress, power delivery, etc.) and optimize design quality (timing, power consumption, etc.). Towards these objectives, the following five projects have been presented in this dissertation:

- A study on the impact of P/G TSV placement styles and TSV RC variation on the 3D power delivery network.
- A full-chip TSV thermo-mechanical stress and reliability analysis flow and optimization methods for TSV-based 3D ICs.
- A chip/package co-analysis of thermo-mechanical stress and reliability as well as mobility and performance variations.
- A study on TSV interfacial crack and substrate crack.
- Physical design methodologies for low power 3D ICs.

The proposed non-regular P/G TSV placement algorithm searches power noisy spots and inserts P/G TSVs iteratively in a greedy fashion. This method significantly reduced the number of P/G TSVs used while satisfying the given IR-drop noise constraint compared

with the conventional regular P/G TSV placement scheme. This in turn helped reduce signal net routing congestions and hence wirelength and footprint area. In addition, the impact of TSV RC variation on the robustness of two-die stacked 3D PDN was examined. Simulation results demonstrated that TSV RC variations cause negligible influence on both static and dynamic noise in 3D PDN due to much smaller RC parasitic values of TSVs compared with that of entire PDN. This indicates that building robust 3D PDN under nominal condition is important. One major limitation of the non-regular T/G TSV placement method is that it only finds P/G TSV locations with given PDN topology and P/G bumps locations. As a follow-up work, an algorithm that optimizes the whole PDN structure concurrently would be worthwhile.

The presented full-chip stress analysis flow provides a quick and reasonably accurate assessment of mechanical reliability problems in TSV-based 3D ICs. This method utilizes the simple but efficient principle of linear superposition of stress tensors, which enables to overcome the limitation of FEA tools. The TSV size and pitch, liner material and its thickness, and TSV placement style were identified as key design knobs to reduce the mechanical reliability problems in TSV-based 3D ICs. One limitation of this linear superposition method is relatively large errors inside TSVs and with smaller TSV pitches. Although, these error conditions are not important for reliability or practical in current TSV technologies, it would be necessary for future TSVs. Another limitation is that all materials in the TSV structure are assumed to be linear elastic. However, if temperature exceeds a certain threshold during thermal cycles, a material will become plastic, not elastic. In this situation, the linear superposition method will no longer hold. Future researches addressing these limitations would be practical and valuable.

The proposed chip/package co-analysis method for thermo-mechanical stress accurately captures the inter-play between chip and package elements within a fraction of runtime of FEA simulations. This flow is built based on the principle of lateral and vertical linear superposition of stress tensors, considering all chip/package elements. It was observed that the mechanical reliability of TSVs in the bottom-most die in the 3D stack are highly affected by packaging elements due to proximity to package-bumps, and that effect decreases

as we go to the upper dies. It was also demonstrated that the carrier mobility variations across the 3D stack and hence full-stack timing variations become non-negligible as both chip and packaging elements are considered. However, this method also assumes that all materials are linear elastic regardless of the temperature. For more accurate and realistic analysis, it would be interesting to incorporate temperature dependent elastic/plastic material properties into this flow.

In the crack study, it was found that the crack propagation behavior for both TSV interfacial crack and substrate crack highly depends on the TSV placement nearby the crack front. It is also worthwhile to note that the TSV liner material and its thickness affect the crack growth significantly by changing the stress magnitude. The proposed DOE and RSM based full-chip TSV interfacial crack analysis flow was utilized to provide design optimization guides for reliability. One limitation of this study is that the impact of crack is confined to the thermo-mechanical reliability only. A comprehensive study considering electrical impact on transistors would be a meaningful follow-up work. Another limitation is that this crack is modeled only for the fabrication process. Continuous thermal cycles during normal chip operation would pose additional stress and hence fatigue around the TSV structure. It would be interesting to study how the chip operation affects the crack growth pattern and reliability.

To maximize power benefit in 3D ICs, the block folding method was found to be a decisive factor. The block folding scheme is circuit specific, e.g., circuit size, wirelength distribution, net power consumption, number/portion of macro blocks, connectivity, number of 3D connections, and so on. In addition, the circuit partitioning is greatly affected by bonding styles largely due to the size and pitch of 3D interconnect elements, i.e., TSV or F2F via. The major limitation of this study is that the block folding scheme was mostly determined manually, although certain criteria were utilized. A systematic approach that concurrently finds optimal circuit partition and TSV/F2F via locations would be very interesting. Although it is possible to perform timing and power optimizations using existing 2D EDA tools with timing constraints on the die boundary ports (= TSVs or F2F vias),

the whole 3D design is not captured by the 2D EDA tools, hence various powerful optimization techniques cannot be performed. Thus, development of a true 3D timing/power optimization engine will help further maximize 3D benefits. In addition, the architecture of both 2D and 3D designs was identical throughout this study. To fully exploit the 3D benefit such as higher bandwidth, a new circuit architecture might be necessary. This architectural optimization for 3D systems would open the whole new degrees of power reduction. Furthermore, this work did not consider manufacturing cost. The 3D power benefit study along with a proper cost modeling will add more value and practicality.

## REFERENCES

- [1] JAEGER, R. C., SUHLING, J. C., RAMANI, R., BRADLEY, A. T., and XU, J., “CMOS Stress Sensors on (100) Silicon,” *IEEE J. Solid-State Circuits*, vol. 35, no. 1, pp. 85–95, 2000.
- [2] JUNG, M., MITRA, J., PAN, D. Z., and LIM, S. K., “TSV Stress-aware Full-Chip Mechanical Reliability Analysis and Optimization for 3D IC,” *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 31, no. 8, pp. 1194–1207, 2012.
- [3] XIONG, W., CLEAVELIN, C. R., KOHLI, P., HUFFMAN, C., SCHULZ, T., SCHRUEFER, K., GEBARA, G., MATHEWS, K., PATRUNO, P., VAILLANT, Y.-M. L., CAYRE-FOURCQ, I., KENNARD, M., MAZURE, C., SHIN, K., and LIU, T.-J. K., “Impact of Strained-Silicon-on-Insulator (sSOI) Substrate on FinFET Mobility,” *IEEE Electron Device Letters*, vol. 27, no. 7, pp. 612–614, 2006.
- [4] YEAP, K. B., ZSCHECHA, E., HANGENA, U. D., WYROBEKA, T., KONGA, L. W., KARMAKARA, A., XUA, X., and PANCHENKO, I., “Elastic anisotropy of Cu and its impact on stress management for 3D IC: Nanoindentation and TCAD simulation study,” *J. Materials Research*, vol. 27, no. 1, pp. 339–348, 2012.
- [5] JUNG, M., MITRA, J., PAN, D. Z., and LIM, S. K., “TSV Stress-aware Full-Chip Mechanical Reliability Analysis and Optimization for 3D IC,” in *Proc. ACM Design Automation Conf.*, pp. 188–193, 2011.
- [6] “International Technology Roadmap for Semiconductors (2012 Update).”
- [7] HUANG, G., BAKIR, M., NAEEMI, A., CHEN, H., and MEINDL, J. D., “Power Delivery for 3D Chips Stacks: Physical Modeling and Design Implication,” in *Proc. IEEE Electrical Performance of Electronic Packaging*, pp. 205–208, 2007.
- [8] HEALY, M. B. and LIM, S. K., “Power Delivery System Architecture for Many-tier 3D Systems,” in *IEEE Electronic Components and Technology Conf.*, pp. 1682–1688, 2010.
- [9] KHAN, N. H., ALAM, S. M., and HASSOUN, S., “System-Level Comparison of Power Delivery Design for 2D and 3D ICs,” in *IEEE Int. 3D System Integration Conf.*, pp. 1–7, 2009.
- [10] YU, H., HO, J., and HE, L., “Simultaneous Power and Thermal Integrity Driven Via Stapling in 3D ICs,” in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 802–808, 2006.
- [11] YANG, J.-S., ATHIKULWONGSE, K., LEE, Y.-J., LIM, S. K., and PAN, D. Z., “TSV Stress Aware Timing Analysis with Applications to 3D-IC Layout Optimization,” in *Proc. ACM Design Automation Conf.*, pp. 803–806, 2010.

- [12] LIU, X., CHEN, Q., DIXIT, P., CHATTERJEE, R., TUMMALA, R. R., and SITARAMAN, S. K., "Failure Mechanisms and Optimum Design for Electroplated Copper Through-Silicon Vias (TSV)," in *IEEE Electronic Components and Technology Conf.*, pp. 624–629, 2009.
- [13] RYU, S.-K., LU, K.-H., ZHANG, X., IM, J.-H., HO, P. S., and HUANG, R., "Impact of Near-Surface Thermal Stresses on Interfacial Reliability of Through-Silicon-Vias for 3-D Interconnects," *IEEE Trans. on Device and Material Reliability*, vol. 11, no. 1, pp. 35–43, 2010.
- [14] LU, K. H., RYU, S.-K., ZHAO, Q., ZHANG, X., IM, J., HUANG, R., and HO, P. S., "Thermal Stress Induced Delamination of Through Silicon Vias in 3-D Interconnects," in *IEEE Electronic Components and Technology Conf.*, pp. 40–45, 2010.
- [15] NAKAMOTO, M., RADOJCIC, R., ZHAO, W., DASARAPU, V. K., KARMARKAR, A. P., and XU, X., "Simulation Methodology and Flow Integration for 3D IC Stress Management," in *Proc. IEEE Custom Integrated Circuits Conf.*, pp. 1–4, 2010.
- [16] BLACK, B., ANNAVARAM, M., BREKELBAUM, N., DEVALE, J., JIANG, L., LOH, G. H., MCCAULEY, D., MORROW, P., NELSON, D. W., PANTUSO, D., REED, P., and RUPLEY, J., "Die Stacking (3D) Microarchitecture," in *Proc. Annual Int. Symp. Microarchitecture*, pp. 469–479, 2006.
- [17] TSAI, Y.-F., WANG, F., XIE, Y., VIJAYKRISHNAN, N., and IRWIN, M. J., "Design Space Exploration for 3-D Cache," *IEEE Trans. on VLSI Systems*, vol. 16, no. 4, pp. 444–455, 2008.
- [18] LEE, Y.-J., HEALY, M., and LIM, S. K., "Co-design of Reliable Signal and Power Interconnects in 3D Stacked ICs," in *Proc. IEEE Int. Interconnect Technology Conference*, pp. 56–58, 2009.
- [19] STARK, D. and HOROWITZ, M., "Techniques for Calculating Currents and Voltages in VLSI Power Supply Networks," *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 9, no. 2, pp. 126–132, 1990.
- [20] TAN, X. D. S. and SHI, C. J. R., "Fast Power/Ground Network Optimization Based on Equivalent Circuit Modeling," in *Proc. ACM Design Automation Conf.*, pp. 550–554, 2001.
- [21] CADENCE, "VoltageStorm." <http://www.cadence.com>.
- [22] SYNOPSYS, "Design Compiler." <http://www.synopsys.com>.
- [23] CADENCE, "Soc Encounter." <http://www.cadence.com>.
- [24] DER PLAS, G. V., LIMAYE, P., MERCHA, A., OPRINS, H., TORREGIANI, C., THIJS, S., LINTEN, D., STUCCHI, M., GURUPRASAD, K., VELENIS, D., SHINICHI, D., CHERMAN, V., VANDEVELDE, B., SIMONS, V., WOLF, I. D., LABIE, R., PERRY, D., BRONCKERS, S., MINAS, N., CUPAC, M., RUYTHOOREN, W., OLMEN, J. V., PHOMMAHAXAY, A., DE POTTER DE TEN BROECK, M., OPDEBEECK, A., RAKOWSKI, M., WACHTER, B. D., DEHAN, M., NELIS, M., AGARWAL, R., DEHAENE, W., TRAVALY, Y., MARCHAL, P., and BEYNE, E., "Design numbers and Considerations for Low-Cost

- 3D TSV IC Technology,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 148–149, 2010.
- [25] TOPOL, A. W., TULIPE, J. D. C. L., SHI, L., FRANK, D. J., BERNSTEIN, K., STEEN, S. E., KUMAR, A., SINGCO, G. U., YOUNG, A. M., GUARINI, K. W., and IEONG, M., “Three-dimensional Integrated Circuits,” *IBM Journal of Research and Development*, vol. 50, no. 4, pp. 35–43, 2006.
  - [26] ENQUIST, P., FOUNTAIN, G., PETTEWAY, C., HOLLINGSWORTH, A., and GRADY, H., “Low Cost of Ownership Scalable Copper Direct Bond Interconnect 3D IC Technology for Three Dimensional Integrated Circuit Applications,” in *IEEE Int. 3D System Integration Conf.*, pp. 1–6, 2009.
  - [27] KATTI, G., STUCCHI, M., MEYER, K. D., and DEHAENE, W., “Electrical Modeling and Characterization of Through Silicon via for Three-Dimensional ICs,” *IEEE Trans. on Electron Devices*, vol. 57, no. 1, pp. 256–262, 2010.
  - [28] SYNOPSYS, “StarRC Parasitic Extraction.”
  - [29] ANSYS, “Q3D Extractor.”
  - [30] SYNOPSYS, “NanoSim.” <http://www.synopsys.com>.
  - [31] SYNOPSYS, “PrimeRail.”
  - [32] “NanGate 45nm Open Cell Library.” <http://http://www.nangate.com/>.
  - [33] LU, K. H., ZHANG, X., RYU, S.-K., IM, J., HUANG, R., and HO, P. S., “Thermo-mechanical Reliability of 3-D ICs Containing Through Silicon Vias,” in *IEEE Electronic Components and Technology Conf.*, pp. 630–634, 2009.
  - [34] LAU, J., ZHENG, X., and SELVANAYAGAM, C., “Failure Analyses of 3D Sip (System-in-Package) and WLP (Wafer-Level Package) by Finite Element Methods,” in *IEEE Int. Symp. on Physical and Failure Analysis of Integrated Circuits*, pp. 108–116, 2009.
  - [35] WU, J. H. and DEL ALAMO, J. A., “Fabrication and Characterization of Through-Substrate Interconnects,” *IEEE Trans. on Electron Devices*, vol. 57, no. 6, pp. 1261–1268, 2010.
  - [36] “FreePDK45.” <http://www.eda.ncsu.edu/wiki/FreePDK>.
  - [37] SIMULIA, “ABAQUS 6.9-EF manual.” <http://www.simulia.com/products/abaqus-cae.html>.
  - [38] ONG, J. M. G., TAY, A. A. O., ZHANG, X., KRIPESH, V., LIM, Y. K., YEO, D., CHEN, K. C., TAN, J. B., HSIA, L. C., and SOHN, D. K., “Optimization of the Thermomechanical Reliability of a 65nm Cu/low-k Large-Die Flip Chip Package,” *IEEE Trans. on Components and Packaging Technology*, vol. 32, no. 4, pp. 838–848, 2009.
  - [39] KARMARKAR, A. P., XU, X., and MOROZ, V., “Performance and Reliability Analysis of 3D-Integration Structures Employing Through Silicon Via (TSV),” in *IEEE Int. Reliability Physics Symposium*, pp. 682–687, 2009.



- [40] YAN, H., TAY, Y. Y., LIANG, M. H., CHEN, Z., NG, C. M., PAN, J. S., XU, H., LIU, C., and SILBERSCHMIDT, V. V., "Amorphous Metallic Thin Films as Copper Diffusion Barrier for Advanced Interconnect Applications," in *IEEE Electronics Packaging Technology Conf.*, pp. 567–572, 2009.
- [41] HOPCROFT, M. A., NIX, W. D., and KENNY, T. W., "What is the Young's Modulus of Silicon," *J. Microelectromechanical Systems*, vol. 19, no. 2, pp. 229–238, 2010.
- [42] McDONOUGH, C., BACKES, B., WANG, W., and GEER, R. E., "Thermal and Spatial Profiling of TSV-induced Stress in 3DICs," in *IEEE Int. Reliability Physics Symposium*, pp. 5D.2.1–5D.2.6, 2011.
- [43] ZHANG, J., BLOOMFIELD, M. O., LU, J.-Q., GUTMANN, R. J., and CALE, T. S., "Modeling Thermal Stresses in 3-D IC Interwafer Interconnects," *IEEE Trans. on Semiconductor Manufacturing*, vol. 19, no. 4, pp. 437–448, 2006.
- [44] XIANG, Y., CHEN, X., and VLASSAK, J. J., "The mechanical properties of electroplated Cu thin films measured by means of the bulge test technique," in *Proc. Material Research Society Symposium*, 2002.
- [45] FRANSSILA, S., "Introduction to Microfabrication." John Wiley and Sons, 2004.
- [46] KIM, D. H., ATHIKULWONGSE, K., and LIM, S. K., "A Study of Through-Silicon-Via Impact on the 3D Stacked IC Layout," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 674–680, 2009.
- [47] YU, L., CHANG, W.-Y., ZUO, K., WANG, J., YU, D., and BONING, D., "Methodology for Analysis of TSV Stress Induced Transistor Variation and Circuit Performance," in *Proc. Int. Symp. on Quality Electronic Design*, pp. 216–222, 2012.
- [48] JUNG, M., LIU, X., SITARAMAN, S. K., PAN, D. Z., and LIM, S. K., "Full-Chip Through-Silicon-Via Interfacial Crack Analysis and Optimization for 3D IC," in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 563–570, 2011.
- [49] ZIMPRICH, P. *et al.*, "Coefficient of Thermal Expansion of Thin Copper Foils for Electronic Laminate Structures," in *Materials Meeting*, 2001.
- [50] PARES, G., BRESSON, N., MINORET, S., LAPRAS, V., BRIANCEAU, P., LUGAND, J. F., ANCIANT, R., and SILLON, N., "Through Silicon Via Technology using Tungsten Metallization," in *Proc. IEEE Int. Conf. on Integrated Circuit Design and Technology*, pp. 1–4, 2011.
- [51] LIU, X., CHEN, Q., SUNDARAM, V., SIMMONS-MATTHEWS, M., WACHTLER, K. P., TUMMALA, R. R., and SITARAMAN, S. K., "Reliability Assessment of Through-Silicon Vias in Multi-Die Stack Packages," *IEEE Trans. on Device and Material Reliability*, vol. 12, no. 2, pp. 263–271, 2012.
- [52] KIM, D. H., TOPALOGLU, R., and LIM, S. K., "Block-level 3D IC Design with Through-Silicon-Via Planning," in *Proc. Asia and South Pacific Design Automation Conf.*, pp. 335–340, 2012.

- [53] ATHIKULWONGSE, K., CHAKRABORTY, A., YANG, J.-S., PAN, D. Z., and LIM, S. K., “Stress-Driven 3D-IC Placement with TSV Keep-Out Zone and Regularity Study,” in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 669–674, 2010.
- [54] LU, K. H., RYU, S.-K., IM, J., HUANG, R., and HO, P. S., “Thermomechanical Reliability of Through-Silicon Vias in 3D Interconnects,” in *IEEE Int. Reliability Physics Symposium*, pp. 3D.1.1–3D.1.7, 2011.
- [55] VEMPATI, S. R., SU, N., KHONG, C. H., LIM, Y. Y., VAIDYANATHAN, K., LAU, J. H., LIEW, B. P., AU, K. Y., TANARY, S., FENNER, A., ERICH, R., and MILLA, J., “Development of 3-D Silicon Die Stacked Package Using Flip Chip Technology with Micro Bump Interconnects,” in *IEEE Electronic Components and Technology Conf.*, pp. 980–987, 2009.
- [56] SYNOPSYS, “PrimeTime.”
- [57] CADENCE, “Encounter Library Characterizer.” <http://www.cadence.com>.
- [58] KIM, J.-S., OH, C. S., LEE, H., LEE, D., HWANG, H.-R., HWANG, S., NA, B., MOON, J., KIM, J.-G., PARK, H., RYU, J.-W., PARK, K., KANG, S.-K., KIM, S.-Y., KIM, H., BANG, J.-M., CHO, H., JANG, M., HAN, C., LEE, J.-B., KYUNG, K., CHOI, J.-S., and JUN, Y.-H., “A 1.2V 12.8GB/s 2Gb Mobile Wide-I/O DRAM with 4x128 I/O Using TSV-Based Stacking,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 496–498, 2011.
- [59] LEE, Y.-J., GOEL, R., and LIM, S. K., “Multi-functional Interconnect Co-optimization for Fast and Reliable 3D Stacked ICs,” in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 645–651, 2009.
- [60] COLETTI, G., TOOL, C., and GEERLIGS, L. J., “Mechanical Strength of Silicon Wafers and its Modelling,” in *Workshop on Crystalline Silicon Solar Cells and Modules: Materials and Processes*, pp. 117–120, 2005.
- [61] ORACLE, “OpenSPARC T2.”
- [62] NAWATHE, U. G., HASSAN, M., WARRINER, L., YEN, K., UPPUTURI, B., GREENHILL, D., KUMAR, A., and PARK, H., “An 8-Core 64-Thread 64b Power-Efficient SPARC SoC,” in *IEEE Int. Solid-State Circuits Conf. Dig. Tech. Papers*, pp. 108–110, 2007.
- [63] KIM, D. H., ATHIKULWONGSE, K., and LIM, S. K., “A Study of Through-Silicon-Via Impact on the 3D Stacked IC Layout,” in *Proc. IEEE Int. Conf. on Computer-Aided Design*, pp. 674–680, 2009.
- [64] LUO, G., SHI, Y., and CONG, J., “An Analytical Placement Framework for 3-D ICs and Its Extension on Thermal Awareness,” *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 4, pp. 510–523, 2013.
- [65] HSU, M.-K., BALABANOV, V., and CHANG, Y.-W., “TSV-Aware Analytical Placement for 3-D IC Designs Based on a Novel Weighted-Average Wirelength Model,” *IEEE Trans. on Computer-Aided Design of Integrated Circuits and Systems*, vol. 32, no. 4, pp. 497–509, 2013.

## PUBLICATIONS

This dissertation is based on and/or related to the works and results presented in the following publications in print:

- [1] **Moongon Jung** and Sung Kyu Lim, “A study of IR-drop noise issues in 3D ICs with Through-Silicon-Vias”, in *IEEE International 3D System Integration Conference*, 2010, pp. 1-7.
- [2] Joydeep Mitra, **Moongon Jung**, Suk-Kyu Ryu, Rui Huang, Sung Kyu Lim, and David Z. Pan, “A Fast Simulation Framework for Full-Chip Thermo-Mechanical Stress and Reliability Analysis of Through-Silicon-Via based 3D ICs”, in *IEEE Electronic Components and Technology Conference*, 2011, pp. 746-753.
- [3] **Moongon Jung**, Joydeep Mitra, David Z. Pan, and Sung Kyu Lim, “TSV Stress-aware Full-Chip Mechanical Reliability Analysis and Optimization for 3D IC”, in *ACM Design Automation Conference*, 2011, pp. 188-193.
- [4] **Moongon Jung**, Shreepad Panth, and Sung Kyu Lim, “A Study of TSV Variation Impact on Power Supply Noise”, in *IEEE International Interconnect Technology Conference*, 2011, pp. 1-3.
- [5] **Moongon Jung**, Xi Liu, Suresh Sitaraman, David Z. Pan and Sung Kyu Lim, “Full-Chip Through-Silicon-Via Interfacial Crack Analysis and Optimization for 3D IC”, in *IEEE/ACM International Conference on Computer-Aided Design*, 2011, pp. 563-570.
- [6] **Moongon Jung** and Sung Kyu Lim, “Chip/Package Co-analysis of Mechanical Reliability in TSV-based 3D ICs”, in *IEEE Workshop on Chip-Packaging Co-Design for High Performance Electronic Systems*, 2011.
- [7] David Z. Pan, Sung Kyu Lim, Krit Athikulwongse, **Moongon Jung**, Joydeep Mitra, Jiwoo Pak, Mohit Pathak, and Jae-seok Yang, “Design for Manufacturability

- and Reliability for TSV-based 3D ICs”, in *IEEE/ACM Asia South Pacific Design Automation Conference*, 2012, pp. 750-755.
- [8] **Moongon Jung**, David Z. Pan, and Sung Kyu Lim, “Chip/Package Co-Analysis of Thermo-Mechanical Stress and Reliability in TSV-based 3D ICs”, in *ACM Design Automation Conference*, 2012, pp. 317-326.
- [9] **Moongon Jung**, Joydeep Mitra, David Z. Pan, and Sung Kyu Lim, “TSV Stress-Aware Full-Chip Mechanical Reliability Analysis and Optimization for 3D IC”, in *SRC Techcon Conference*, 2012.
- [10] **Moongon Jung**, Joydeep Mitra, David Z. Pan, and Sung Kyu Lim, “TSV Stress-aware Full-Chip Mechanical Reliability Analysis and Optimization for 3D IC”, in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 31, No. 8, pp. 1194-1207, 2012.
- [11] **Moongon Jung**, Christoph Sander, Uwe Muehle, KongBoon Yeap, David Z. Pan, and Sung Kyu Lim, “Impact of Material Property Variations on Full-Chip Reliability and Performance in TSV-based 3D ICs”, in *SRC Techcon Conference*, 2013.
- [12] **Moongon Jung**, Young-Joon Lee, Taigon Song, Yang Wan, and Sung Kyu Lim, “Design Methodologies for Low Power 3D Processors”, in *SRC Techcon Conference*, 2013.
- [13] Chun Zhang, **Moongon Jung**, Sung Kyu Lim and Yiyu Shi, “Novel Crack Sensor Design for TSV-based 3D Integrated Circuits: Design and Deployment Perspectives”, in *IEEE/ACM International Conference on Computer-Aided Design*, 2013, pp. 371-378.
- [14] **Moongon Jung**, Taigon Song, Yang Wan, Young-Joon Lee, Debabrata Mohapatra, Hong Wang, Greg Taylor, Devang Jariwala, Vijay Pitchumani, Patrick Morrow, Clair Webb, Paul Fischer, and Sung Kyu Lim, “How to Reduce Power in 3D IC Designs: A Case Study with OpenSPARC T2 Core”, in *IEEE Custom Integrated Circuits Conference*, 2013, pp. 1-4.

- [15] **Moongon Jung**, David Z. Pan, and Sung Kyu Lim, “Chip/Package Mechanical Stress Impact on 3D IC Reliability and Mobility Variations”, in *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, Vol. 32, No. 11, pp. 1694-1707, 2013.
- [16] **Moongon Jung**, Joydeep Mitra, David Z. Pan, and Sung Kyu Lim, “Full-Chip Mechanical Reliability Analysis and Optimization for 3D ICs”, in *Communications of the ACM, Research Highlight*, Vol. 57, No. 1, pp. 107-115, 2014.
- [17] **Moongon Jung**, Taigon Song, Yang Wan, Yarui Peng, and Sung Kyu Lim, “On Enhancing Power Benefits in 3D ICs: Block Folding and Bonding Styles Perspective”, in *ACM Design Automation Conference*, 2014, to appear.

In addition, the author has completed works unrelated to this dissertation presented in the following publications in print:

- [1] Young-Joon Lee, Mohit Pathak, Chang Liu, **Moongon Jung**, and Sung Kyu Lim, “Design and Timing Optimization of a 3D Stacked Microprocessor”, in *ACM International Workshop on Timing Issues in the Specification and Synthesis of Digital Systems*, 2010.
- [2] Michael B. Healy, Krit Athikulwongse, Rohan Goel, Mohammad M. Hossain, Dae Hyun Kim, Young-Joon Lee, Dean L. Lewis, Tzu-Wei Lin, Chang Liu, **Moongon Jung**, Brian Ouellette, Mohit Pathak, Hemant Sane, Guanhao Shen, Dong Hyuk Woo, Xin Zhao, Gabriel H. Loh, Hsien-Hsin S. Lee, and Sung Kyu Lim, “Design and Analysis of 3D-MAPS: A Many-Core 3D Processor with Stacked Memory”, in *IEEE Custom Integrated Circuits Conference*, 2010, pp. 1-4.
- [3] Dae Hyun Kim, Krit Athikulwongse, Michael B. Healy, Mohammad M. Hossain, **Moongon Jung**, Ilya Khorosh, Gokul Kumar, Young-Joon Lee, Dean L. Lewis, Tzu-Wei Lin, Chang Liu, Shreepad Panth, Mohit Pathak, Minzhen Ren, Guanhao Shen, Taigon Song, Dong Hyuk Woo, Xin Zhao, Joungho Kim, Ho Choi, Gabriel H. Loh, Hsien-Hsin S. Lee, and Sung Kyu Lim, “3D-MAPS: 3D Massively Parallel Processor

with Stacked Memory”, in *IEEE International Solid-State Circuits Conference*, 2012, pp. 188-190.

- [4] Krit Athikulwongse, Dae Hyun Kim, **Moongon Jung**, Sung Kyu Lim, “Block-level Designs of Die-to-Wafer Bonded 3D ICs and Their Design Quality Tradeoffs”, in *IEEE/ACM Asia South Pacific Design Automation Conference*, 2013, pp. 687-692.

## VITA

Moongon Jung was born in Pusan, Republic of Korea, in 1980. He received the B.S. degree from Seoul National University, Seoul, Korea, in 2003, and the M.S. degree from Stanford University, Stanford, CA, in 2009, both in electrical engineering. He is currently a Ph.D candidate in the School of Electrical and Computer Engineering at Georgia Institute of Technology.

From 2009 to present, he has been a graduate research assistant in Georgia Tech Computer-Aided Design (GTCAD) laboratory led by Professor Sung kyu Lim. He has been working in the areas of physical design methods for low power 3D ICs, thermo-mechanical reliability analysis and optimization of TSV-based 3D ICs, and 3D power delivery network analysis and optimization.

His works were nominated for the Best Paper Award at ACM Design Automation Conference (DAC) 2011 and 2012 and IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems (TCAD) 2014.